

STIC-Biotech/ChemLib

142525

From: Priebe, Scott
Sent: Thursday, January 13, 2005 9:24 AM
To: STIC-Biotech/ChemLib
Cc: Whiteman, Brian
Subject: FW: sequence search - RUSH

Importance: High

Please RUSH the requested search.

-----Original Message-----

From: Whiteman, Brian
Sent: Thursday, January 13, 2005 9:10 AM
To: STIC-Biotech/ChemLib; Priebe, Scott
Subject: sequence search

10/057,136
EFD 2/24/97
Inventor: Schlom et al.

Please perform an oligonucleotide search for SEQ ID NOs: 2 and 4-12 against public databases, us patent, and us patent application databases.

Thank you,

Brian Whiteman
Remsen, 2D14
mail box 2C18
Patent Examiner - Art Unit 1635
United States Patent and Trademark Office
(571) 272-0764

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2- _____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA Sequence: # _____
AA Sequence :# _____
Structure: # _____
Bibliographic: _____
Litigation: _____
Patent Family: _____
Other: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)

Query Match	100.0%	Score 60;	DB 14;	Length 60;
Best Local Similarity	100.0%	Pred. No. 5.2e-13;		
Matches 60;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GCTAGTACAGCCGACCCGACATGGCGGTACGAGCGCTCCGATACGAGACCGGGCCT	60	

Db 1 GGTAAGTACAGCGCCGACATGGCGTCAAGAGCGCTCCGGATACGAGACCGGCGCT 60

RESULT 2

US-10-406-317-41
; Sequence 41, Application US/10406317
; Publication No. US20040019195A1
; GENERAL INFORMATION:
; APPLICANT: Schlom, Jeffrey;
; APPLICANT: Hodge, James;
; APPLICANT: Panicali, Dennis
; TITLE OF INVENTION: A recombinant vector expressing multiple constitutulatory
; FILE REFERENCE: 38163-0189
; CURRENT APPLICATION NUMBER: US/10/406,317
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US/09/856,988
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: PCT/US99/26866
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/111,582
; PRIOR FILING DATE: 1998-12-09
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 41
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: VECTOR
US-10-406-317-41

Query Match 100.0%; Score 60; DB 16; Length 2297;
Best Local Similarity 100.0%; Pred. No. 4.5e-13;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTAAGTACAGCGCCGACATGGCGTCAAGAGCGCTCCGGATACGAGACCGGCGCT 60
Db 886 GGTAAGTACAGCGCCGACATGGCGTCAAGAGCGCTCCGGATACGAGACCGGCGCT 945

RESULT 3

US-10-057-136-9
; Sequence 9, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 9
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-9

Query Match 63.7%; Score 38.2; DB 14; Length 60;
Best Local Similarity 78.0%; Pred. No. 8.6e-05;

Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGTAAGTACAGCGCCGACATGGCGTCAAGAGCGCTCCGGATACGAGACCGGCGCC 59
Db 1 GGAAGTACCGCTCCACTGTCACACGCGGCTCACAAGCGCCGACAGACTGACTGCGCC 59

RESULT 4

US-10-057-136-4
; Sequence 4, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 4
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-4

Query Match 62.7%; Score 37.6; DB 14; Length 60;
Best Local Similarity 76.7%; Pred. No. 0.00015;
Matches 46; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 GGTAAGTACAGCGCCGACATGGCGTCAAGAGCGCTCCGGATACGAGACCGGCGCT 60
Db 1 GGAAGTACTGCACCGCACCATGGCGTCAAGAGCGCTCCGGATACGAGACCGGCGCT 60

RESULT 5

US-10-057-136-8
; Sequence 8, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 8
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-8

Query Match 61.0%; Score 36.6; DB 14; Length 60;
Best Local Similarity 76.3%; Pred. No. 0.00035;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 GGTAGTACAGCGCCGACATGGGCTCAGAGCGCTCCGATACGAGACCGCGCC 59
DB 1 GGCAGCAGCCGACCGCCGACAGCGGTCAAGCGCGCAGACACTGACCTGCGCC 59

RESULT 6

US-10-057-136-7
; Sequence 7, Application US/10057136
; Publication No. US20030021770A1

; GENERAL INFORMATION:

; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA

; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C

; CURRENT APPLICATION NUMBER: US/10/057, 136
; CURRENT FILING DATE: 2002-01-25

; PRIOR APPLICATION NUMBER: 09/366, 670
; PRIOR FILING DATE: 1999-08-03

; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24

; PRIOR APPLICATION NUMBER: 60/038, 253
; PRIOR FILING DATE: 1997-02-24

; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 7

; LENGTH: 60

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-057-136-7

Query Match 54.7%; Score 32.8; DB 14; Length 60;
Best Local Similarity 71.7%; Pred. No. 0.0094;
Matches 43; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 GGTAGTACAGCGCCGACATGGGCTCAGAGCGCTCCGATACGAGACCGCGCC 60
DB 1 GGTTCGACGCGCCCGCTCTGCTACGGTGTACATCCGCGCGGATACGAGACCGCGCCCT 60

RESULT 7

US-10-057-136-6

; Sequence 6, Application US/10057136
; Publication No. US20030021770A1

; GENERAL INFORMATION:

; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA

; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C

; CURRENT APPLICATION NUMBER: US/10/057, 136
; CURRENT FILING DATE: 2002-01-25

; PRIOR APPLICATION NUMBER: 09/366, 670
; PRIOR FILING DATE: 1999-08-03

; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24

; PRIOR APPLICATION NUMBER: 60/038, 253
; PRIOR FILING DATE: 1997-02-24

; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6

; LENGTH: 60

; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-6

Query Match 48.3%; Score 29; DB 14; Length 60;
Best Local Similarity 71.7%; Pred. No. 0.26;
Matches 38; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 7 ACAGCCGACCGCCGACATGGGCTCAGAGCGCTCCGATACGAGACCGCGCC 59
DB 7 ACAGCTCTCCCGCTCATGGGTTACTTCTGTGCTCCAGATACTCGCCAGCTCC 59

RESULT 8

US-10-296-734-1165

; Sequence 1165, Application US/10296734
; Publication No. US20040054137A1

; GENERAL INFORMATION:

; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A

; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine

; CURRENT APPLICATION NUMBER: US/10/296, 734
; CURRENT FILING DATE: 2003-08-04

; PRIOR APPLICATION NUMBER: AU P07761/00
; PRIOR FILING DATE: 2000-05-26

; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 1165

; LENGTH: 72

; TYPE: DNA

; ORGANISM: Artificial

; FEATURE:

; OTHER INFORMATION: MUC1F segment 8

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(72)

US-10-296-734-1165

Query Match 48.0%; Score 28.8; DB 16; Length 72;
Best Local Similarity 69.6%; Pred. No. 0.3;
Matches 39; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 GGTAGTACAGCGCCGACATGGGCTCAGAGCGCTCCGATACGAGACCGCGC 56
DB 16 GGAAGCACACCCCTCCCGCTCAGCATGTGACAGCGCTCCCGATACAAAGCCGC 71

RESULT 9

US-10-296-734-1209

; Sequence 1209, Application US/10296734
; Publication No. US20040054137A1

; GENERAL INFORMATION:

; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor.
; FILE REFERENCE: Savine

; CURRENT APPLICATION NUMBER: US/10/296, 734
; CURRENT FILING DATE: 2003-08-04

; PRIOR APPLICATION NUMBER: AU P07761/00
; PRIOR FILING DATE: 2000-05-26

; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 1209

; LENGTH: 16638

; TYPE: DNA

; ORGANISM: Artificial

; FEATURE:

; OTHER INFORMATION: Differentiation svaine

; NAME/KEY: CDS

; LOCATION: (1)..(16638)

US-10-296-734-1209

Query Match 48.0%; Score 28.8; DB 16; Length 16638;
Best Local Similarity 69.6%; Pred. No. 0.24;
Matches 39; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 GGTAGTACAGCCCGCACATGGCGTACAGAGCGCTCCGATACGAGACCGGC 56
DB 4246 GGAAGACACACCTCTCCGCTCAGATGTGACAGCGCTCCGATACGAGACCGGC 4301

RESULT 10

US-10-057-136-2
; Sequence 2, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-2

Query Match 45.7%; Score 27.4; DB 14; Length 60;
Best Local Similarity 69.8%; Pred. No. 1;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACAGCGCCACCGCACATGGCGTACAGAGCGCTCCGATACGAGACCGCGCC 59
DB 7 ACCGCCCCCGCACCGCACGCTGTACCTCGCGCCCGACACAGCGCGCGCC 59

RESULT 11

US-10-447-839A-75/c
; Sequence 75, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald W.
; APPLICANT: Kharbada, Surender
; APPLICANT: Weitman, Steven D.
; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT APPLICATION NUMBER: US/10/447,839A
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 75
; LENGTH: 1424
; TYPE: RNA
; ORGANISM: ARTIFICIAL
; FEATURE:

; OTHER INFORMATION: Synthesized Sequence
US-10-447-839A-75

Query Match 45.7%; Score 27.4; DB 16; Length 1424;
Best Local Similarity 69.8%; Pred. No. 0.9;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACAGCGCCACCGCACATGGCGTACAGAGCGCTCCGATACGAGACCGCGCC 59
DB 1035 ACCGCCCCCGCACCGCACGCTGTACCTCGCGCCCGACACAGCGCGCC 983

RESULT 12

US-10-447-839A-20
; Sequence 20, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald W.
; APPLICANT: Kharbada, Surender
; APPLICANT: Weitman, Steven D.
; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT APPLICATION NUMBER: US/10/447,839A
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 1428
; TYPE: RNA
; ORGANISM: RNA
US-10-447-839A-20

Query Match 45.7%; Score 27.4; DB 16; Length 1428;
Best Local Similarity 67.9%; Pred. No. 0.9;
Matches 36; Conservative 1; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACAGCGCCACCGCACATGGCGTACAGAGCGCTCCGATACGAGACCGCGCC 59
DB 391 ACCGCCCCCGCACCGCACGCTGTACCTCGCGCCCGACACAGCGCGCC 443

RESULT 13

US-10-057-136-19
; Sequence 19, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 1527
; TYPE: DNA

This Page Blank (uspto)

Query Match 45.7%; Score 27.4; DB 4; Length 1804;
Best Local Similarity 69.8%; Pred. No. 1.2;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 7 ACAGCGCCACCCGACATGGCGTCAAGAGCGCTCCGATACGAGACCGCGCC 59
Db 463 ACCGCCCCCAGCCGACCGGTGTCACTCGGCCCCGAGACCAAGCGCGCCCC 515

RESULT 3

US-10-029-517-19
; Sequence 19, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 19
; LENGTH: 8186
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 6899
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7155
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7184
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7957
; OTHER INFORMATION: unknown
; NAME/KEY: intron
; LOCATION: (2997)...(3498)
; OTHER INFORMATION: intron 1
; NAME/KEY: intron:exon junction
; LOCATION: (3498)...(3499)
; OTHER INFORMATION: intron 1:exon 2
; NAME/KEY: exon
; LOCATION: (3508)...(3599)
; OTHER INFORMATION: exon 2d
; NAME/KEY: exon:intron junction
; LOCATION: (3982)...(3983)
; OTHER INFORMATION: exon 2a:intron 2a
; NAME/KEY: intron:exon junction
; LOCATION: (4205)...(4206)
; OTHER INFORMATION: intron 2c:exon 3c
; NAME/KEY: intron:exon junction
; LOCATION: (4259)...(4260)
; OTHER INFORMATION: intron 2d:exon 3d
; NAME/KEY: exon
; LOCATION: (4260)...(4328)
; OTHER INFORMATION: exon 3d
; NAME/KEY: intron:exon junction
; LOCATION: (4632)...(4633)
; OTHER INFORMATION: intron 3:exon 4
; NAME/KEY: exon
; LOCATION: (4914)...(5035)
; OTHER INFORMATION: exon 5
; NAME/KEY: intron
; LOCATION: (5266)...(6293)
; OTHER INFORMATION: intron 6
US-10-029-517-19

Query Match 45.7%; Score 27.4; DB 4; Length 8186;
Best Local Similarity 69.8%; Pred. No. 1.4;

Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 7 ACAGCGCCACCCGACATGGCGTCAAGAGCGCTCCGATACGAGACCGCGCC 59
Db 3831 ACCGCCCCCAGCCGACCGGTGTCACTCGGCCCCGAGACCAAGCGCGCCC 3883

RESULT 4

US-09-860-473-28/c
; Sequence 28, Application US/09860473
; Patent No. 6656732
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF SRC-C EXPRESSION
; FILE REFERENCE: RTS-0222
; CURRENT APPLICATION NUMBER: US/09/860,473
; CURRENT FILING DATE: 2001-05-18
; NUMBER OF SEQ ID NOS: 169
; SEQ ID NO 28
; LENGTH: 336
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-860-473-28

Query Match 43.7%; Score 26.2; DB 4; Length 336;
Best Local Similarity 67.3%; Pred. No. 2.7;
Matches 37; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGTAGTACAGCGCCACCCGACATGGCGTCAAGAGCGCTCCGATACGAGACCGG 55
Db 107 GGACGTCCAGCGGAGACGGGCAAGCGCGCTGCGCTCCGCTCCAGGCGCG 53

RESULT 5

US-10-029-517-101
; Sequence 101, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 101
; LENGTH: 518
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-10-029-517-101

Query Match 43.3%; Score 26; DB 4; Length 518;
Best Local Similarity 70.0%; Pred. No. 3.3;
Matches 35; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 7 ACAGCGCCACCCGACATGGCGTCAAGAGCGCTCCGATACGAGACCGGC 56
Db 468 ACCGCCCCCAGCCGACCGGTGTCACTCGGCCCCGAGACCAAGCGCGGC 517

RESULT 6

US-10-029-517-18
; Sequence 18, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517

; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 2874
; LENGTH: 2397
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-2874

Query Match 40.0%; Score 24; DB 4; Length 2397;
Best Local Similarity 64.3%; Pred. No. 18;
Matches 36; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 5 GTACAGCGCCACCCGACATGCGCTCAGCAGCGCTCCGGATACGAGACCGGCCT 60
DB 1083 GGACAGCGCTACCCCGCAAATGACAGCGGTGAACCTTAACGATACCGACACCT 1138

RESULT 12

US-09-105-537-40
; Sequence 40, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438Ust
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 40
; LENGTH: 2787
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-105-537-40

Query Match 39.3%; Score 23.6; DB 3; Length 2787;
Best Local Similarity 69.6%; Pred. No. 25;
Matches 32; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 15 ACCCGACATGGCGTCACGAGCGCTCCGGATACGAGACCGGCCT 60
DB 1292 ACCCGACATGACGACCGCGGCTCTTGCACCTTTCGACCGGCTCCT 1337

RESULT 13

US-09-320-878-21
; Sequence 21, Application US/09320878A
; Patent No. 6117659
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/320,878A
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: CIP OF 09/141,908
; EARLIER FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: 60/119,139
; EARLIER FILING DATE: 1999-02-08
; EARLIER APPLICATION NUMBER: 60/100,880

; EARLIER FILING DATE: 1998-09-22
; EARLIER APPLICATION NUMBER: 60/087,080
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 21
; LENGTH: 5970
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-320-878-21

Query Match 39.3%; Score 23.6; DB 3; Length 5970;
Best Local Similarity 69.6%; Pred. No. 27;
Matches 32; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 15 ACCCGACATGGCGTCACGAGCGCTCCGGATACGAGACCGGCCT 60
DB 4030 ACCCGACATGACGACCGCGGCTCTTGCACCTTTCGACCGGCTCCT 4075

RESULT 14

US-09-141-908-11
; Sequence 11, Application US/09141908
; Patent No. 6503741
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: Combinatorial Polyketide Libraries Produced Using a
; FILE REFERENCE: 300622002100
; CURRENT APPLICATION NUMBER: US/09/141,908
; CURRENT FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: PROV. 60/076,919
; EARLIER FILING DATE: 1998-03-05
; EARLIER APPLICATION NUMBER: PROV. 60/087,080
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 11
; LENGTH: 5970
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-141-908-11

Query Match 39.3%; Score 23.6; DB 4; Length 5970;
Best Local Similarity 69.6%; Pred. No. 27;
Matches 32; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 15 ACCCGACATGGCGTCACGAGCGCTCCGGATACGAGACCGGCCT 60
DB 4030 ACCCGACATGACGACCGCGGCTCTTGCACCTTTCGACCGGCTCCT 4075

RESULT 15

US-09-657-440-21
; Sequence 21, Application US/09657440
; Patent No. 6509455
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/657,440

; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 09/320,878
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: CIP OF 09/141,908
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 5970
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-657-440-21

Query Match 39.3%; Score 23.6; DB 4; Length 5970;
Best Local Similarity 69.6%; Pred. No. 27;
Matches 32; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 15 ACCCGCACATGGCGTCACGAGCGCTCCGATACGAGACCGGCGCT 60
|||||
Db 4030 ACCCGCACATGACGAGCGGCGGCGCTCCGACTCTTCGACCGGCTCCT 4075
|||||

Search completed: January 16, 2005, 03:02:00
Job time : 44.2 secs

This Page Blank (uspto)

QY 7 ACAGCGCCACCCGACATGCGCTCAGCAGCGCTCCGGATACGAGACCGCGGCC 59
Db 250 ACCGCCCCCCCCAGCGCTGTCTCAGCTCGGCTCCGGACACCAAGCGCGGCC 198

RESULT 2
LOCUS BU152566
DEFINITION BU152566 1349 bp mRNA linear EST 03-SEP-2002
AGENCOURT_8670683 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6380559
5', mRNA sequence.
BU152566
BU152566.1 GI:22666098
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
1 (bases 1 to 1349)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LLCM2569 row: f column: 16
High quality sequence stop: 291.
Location/Qualifiers

FEATURES
source
1..1349
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6380559"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 40"
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 48.3%; Score 29; DB 5; Length 1349;
Best Local Similarity 71.7%; Pred. No. 18;
Matches 38; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 7 ACAGCGCCACCCGACATGCGCTCAGCAGCGCTCCGGATACGAGACCGCGGCC 59
Db 265 ACCGCCCCCCCCAGCGCTGTCTCAGCTCGGCTCCGGACACCAAGCGCGGCC 317

RESULT 3
LOCUS BY256331/c 486 bp mRNA linear EST 10-DEC-2002
DEFINITION BY256331 RIKEN full-length enriched, visual cortex Mus musculus
cDNA clone K330023N10 5', mRNA sequence.
BY256331
BY256331.1 GI:26437843
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 486)

REFERENCE

AUTHORS
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,
Choctia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
Gustringich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S.,
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
Sandelin, A., Schneider, C., Sempke, C.A., Setou, M., Shimada, K.,
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
22354683
12466851
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.jp, <http://genome.gsc.riken.jp/>
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,
Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,
Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,
Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,
Shiraki, T., Tagami, M., Waki, K., Watanabe, A., Muramatsu, M. and
Hayashizaki, Y. Direct Submision
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Michela Fagioli and Takao K. Hensch (
Laboratory for Neuronal Circuit Development Brain Science Institute
RIKEN 2-1 Hirosawa, Wako-shi, Saitama 351-0198 Japan) whose
assistance we gratefully acknowledge. Please visit our web site
(<http://genome.gsc.riken.go.jp>) for further details.
Location/Qualifiers
1..486
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"

FEATURES
source

ORIGIN

/clone="K330023N10"
/tissue_type="visual cortex"
/clone_lib="RIKEN full-length enriched, visual cortex"

Query Match 46.0%; Score 27.6; DB 5; Length 486;
Best Local Similarity 66.1%; Pred. No. 57;
Matches 39; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 2 GTAGTACAGCGCCACCGCATGCGCTCAGCAGCGCTCCGATACGAGACCGCGCCT 60
123 GTAGCGCGCGCGCGCCCTTACGTGCTTCACGTGCGCAGGAAAAACAGAACCCGCGCCT 65

RESULT 4
LOCUS BM791359 604 bp mRNA linear EST 05-MAR-2002
DEFINITION K-EST0071342 S21SNU520 Homo sapiens cDNA clone S21SNU520-14-A06 5',
BM791359 mRNA sequence.
ACCESSION BM791359.1 GI:19139591
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 604)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 14 row: A column: 06
High quality sequence stop: 604.

FEATURES

source

1..604
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S21SNU520-14-A06"
/sex="F"
/tissue_type="Stomach"
/cell_type="Floating aggregates"
/cell_line="SNU-520"
/lab_host="Top10F"
/clone_lib="S21SNU520"
/note="Organ: Stomach; Vector: pTZ18RPI; Site_1: EcoRI;
Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN

Query Match 45.7%; Score 27.4; DB 4; Length 604;
Best Local Similarity 69.8%; Pred. No. 65;

Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 7 ACAGCGCCACCGCAGCATGCGCTACGAGCGCTCCGATACGAGACCGCGCC 59
34 ACCGACCCCGCAGCCACGCTGTACCTCGGCCCCGACACGAGCGCGCCCC 86

RESULT 5
LOCUS BU542454 877 bp mRNA linear EST 13-SEP-2002
DEFINITION AGENCOURT_10322173 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574322
5', mRNA sequence.
ACCESSION BU542454
VERSION BU542454.1 GI:22852937
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens (human)
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 877)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM2769 row: h column: 02
High quality sequence stop: 760.

FEATURES

source

1..877
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6574322"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 40"
/note="Organ: prostate; Vector: POTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACGAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 45.7%; Score 27.4; DB 5; Length 877;
Best Local Similarity 69.8%; Pred. No. 64;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACAGCGCCACCGCAGCATGCGCTACGAGCGCTCCGATACGAGACCGCGCC 59
41 ACCGCCCCCGCAGCCACGCTGTACCTCGGCCCCGACACGAGCGCGCCCC 93

RESULT 6
LOCUS CA489836 959 bp mRNA linear EST 14-NOV-2002
DEFINITION AGENCOURT_10810668 MAPcL Homo sapiens cDNA clone IMAGE:6722324 5',
CA489836 mRNA sequence.
ACCESSION CA489836
VERSION CA489836.1 GI:24952627
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 959)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/ntp

CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>

Plate: L1AM14284 row: n column: 20
High quality sequence start: 31
High quality sequence stop: 446.
Location/Qualifiers

FEATURES
source

1. 959

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6722324"

/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HMEL, LNCaP"

/lab_host="EMD10B"

/clone_lib="MAPCL"

/note="Vector: pCMV-SPORT6; Site_1: EcoRV; Site_2: Not I; Subtracted with brain, liver, lung, kidney and muscle. Directionally cloned. Priming method: oligo-dT. Average insert size: 1800 bp. Library amplification: 26,000 fold. Kristi A. Egland, James J. Vincent, Robert Strausberg, Bungkok Lee & Ira Pastan. Discovery of new breast cancer genes encoding membrane and secreted proteins. Manuscript submitted."

ORIGIN

Query Match 45.7%; Score 27.4; DB 6; Length 959;
Best Local Similarity 69.8%; Pred. No. 64;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACAGCGCCACCCGACATGCGTCACGAGCGCTCCGGATACGAGACCGCGCC 59
Db 355 ACCGCCCCCAGCCGCGGTTCACCTCGCGCCCGGACACCGCGCGCC 407

RESULT 7
BU148487
LOCUS

DEFINITION BU148487 1113 bp mRNA linear EST 03-SEP-2002
AGENCOURT_8670479 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6380642
5', mRNA sequence.

ACCESSION BU148487
VERSION BU148487.1 GI:22662019
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1113)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/ntp

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: L1CM2569 row: j column: 03
High quality sequence stop: 235.

FEATURES
source

Location/Qualifiers

1. 1113

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6380642"

/tissue_type="carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_40"

/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 45.7%; Score 27.4; DB 5; Length 1113;
Best Local Similarity 69.8%; Pred. No. 63;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACAGCGCCACCCGACATGCGTCACGAGCGCTCCGGATACGAGACCGCGCC 59
Db 36 ACCGCCCCCAGCCGCGGTTCACCTCTGCCCGGACACCGCGCGCC 88

RESULT 8
BU542996
LOCUS

DEFINITION BU542996 1130 bp mRNA linear EST 13-SEP-2002
AGENCOURT_10338707 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574979
5', mRNA sequence.

ACCESSION BU542996
VERSION BU542996.1 GI:22853479
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1130)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/ntp

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: L1CM2771 row: c column: 11
High quality sequence start: 27
High quality sequence stop: 246.
Location/Qualifiers

FEATURES
source

1. 1130

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6574979"

/tissue_type="carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_40"

/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 45.7%; Score 27.4; DB 5; Length 1130;
Best Local Similarity 69.8%; Pred. No. 63;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACAGCGCCACCCGACATGCGCTCAGCAGCGCTCCGATACGAGACCGCGCC 59
DB 75 ACCGCCCCCAGCCCGGTGTCTACCTCGGCCCGGACACCAAGCGCGCCC 127

RESULT 9

BO936898 1234 bp mRNA linear EST 21-AUG-2002
LOCUS AGENCOURT_8919110 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6383855
DEFINITION 5', mRNA sequence.

ACCESSION BO936898
VERSION BO936898.1 GI:22352281

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DRP

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2577 row: 0 column: 24
High quality sequence stop: 245.

FEATURES

source

1..1234
Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6383855"

/tissue_type="carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_1lb="NIH_MGC_40"

/note="Organ: prostate; Vector: pOTB7, Site_1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 45.7%; Score 27.4; DB 5; Length 1234;
Best Local Similarity 69.8%; Pred. No. 63;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACAGCGCCACCCGACATGCGCTCAGCAGCGCTCCGATACGAGACCGCGCC 59
DB 36 ACCGCCCCCAGCCCGGTGTCTACCTCGGCCCGGACACCAAGCGCGCCC 88

RESULT 10

BO943554 1268 bp mRNA linear EST 21-AUG-2002
LOCUS AGENCOURT_8777815 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6384308
DEFINITION 5', mRNA sequence.
ACCESSION BO943554
VERSION BO943554.1 GI:22359032

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DRP

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2579 row: b column: 21
High quality sequence stop: 177.

FEATURES

source

1..1268
Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6384308"

/tissue_type="carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_1lb="NIH_MGC_40"

/note="Organ: prostate; Vector: pOTB7, Site_1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 45.7%; Score 27.4; DB 5; Length 1268;
Best Local Similarity 69.8%; Pred. No. 63;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACAGCGCCACCCGACATGCGCTCAGCAGCGCTCCGATACGAGACCGCGCC 59
DB 96 ACCGCCCCCAGCCCGGTGTCTACCTCGGCCCGGACACCAAGCGCGCCC 148

RESULT 11

BO920055 1343 bp mRNA linear EST 20-AUG-2002
LOCUS AGENCOURT_10031674 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6481609
DEFINITION 5', mRNA sequence.

ACCESSION BO920055
VERSION BO920055.1 GI:22334753

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DRP

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: LLCM2663 row: e column: 02
High quality sequence start: 56
High quality sequence stop: 237.

FEATURES

Source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6481609"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
Note: this is a NIH MGC Library."

```

ORIGIN

Query Match	45.7%	Score 27.4;	DB 5;	length 1343;
Best Local Similarity	69.8%;	Pred. No. 63;		
Matches 37; Conservative	0;	Mismatches 16;	Indels 0	2

```

QY      7 ACAGCGGCACCCGCAATGGCGTCAAGCGCTCCGGATACGAGCCGGCGCC 59
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      78 ACCGCCCCCGCAGCCCAAGTGTCACTCTGCCCCCGACACGAGCCGGCCCC 130

```

RESULT 12	
BUS42790	
LOCUS	
DEFINITION	BUS42790
ACCESSION	AGENCOCURT_10334841 NIH_MGC_40 Homo sapiens CDNA clone IMAGE:6574725 5', mRNA Sequence.

ACCESSION	EU542790	GenBank accession.
VERSION	EU542790.1	
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1 (bases 1 to 1420)	

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 1420)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg ph n

FEATURES

Source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6574725"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit

```

(Stratagene) and Superscript II RT (Life Technologies)
Note: this is a NIH_MGC Library. "

ORIGIN

Query Match	45.7%;	Score 27.4;	DB 5;	Length 1420;
Best Local Similarity	69.8%;	Pred. No. 63;		
Matches 37; Conservative	0;	Mismatches 16;	Indels 0;	Gaps 0;

QY 7 ACAGGGCCACCCGCACATGGCGTCAAGAGCGTCCGGATACGAGACCGCGCC 59
Db 25 ACCGCCCCCCCAAGCCACGAGGTACCTCGGCCCCCGACACCAAGCGCGCCCC 77

RESULT 13	
BUS43309	
LOCUS	BUS43309
DEFINITION	AGENCOURT_10327072 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6575349
	5' mRNA sequence.
	1531 bp
	mRNA linear EST 13-SEP-2002

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE	1 (Bases 1 to 1531)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg ph n

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM2772 row: b column: 21
High quality sequence stop: 166.

FEATURES

Source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6575349"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1b="N1H_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)".
Note: this is a N1H_MGC Library."

```

ORIGIN

Query Match	45.7%;	Score 27.4;	DB 5;	Length 1531;
Best Local Similarity	69.8%;	Pred. No. 62;		
Matches 37; Conservative	0;	Mismatches 16;	Indels 0;	Cov 0

Ox 7 ACAGCGCACCCGACATGGCGTCAGAGCGCTCCGATACGAGCCGCC 59
 | | | | | | | | | | | | | | | | | | | |
Db 36 ACCGCCCCCAGCCCAAGGTGTCACTTCTGCCCGGACACAGGCCGCC 88
 | | | | | | | | | | | | | | | | | | | |

RESULT 14	W16262/c	LOCUS	DEFINITION
W16262	451 bp	mRNA	linear
mb64c07.r1	Soares mouse	p3NMf19.5	Mus musculus CDNA clone
IMAGE:334188	5' similar to gb:Z28407	60S RIBOSOMAL	PROTEIN L8
			EST 12-SEP-1996

(HUMAN); mRNA sequence.
 W16262
 VERSION W16262.1 GI:1537450
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 451)
 Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1290651.
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
 MGI:215588
 Seq primer: -28M13 rev2 from Amerham
 High quality sequence stop: 403.
 FEATURES
 source
 1..451
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:334188"
 /dev_stage="19.5 dpc total fetus"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares mouse p3NM19.5"
 /note="Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dt) primer [5', TGTTACCAATCTGAAGTGGAGCGCGCGCATTTTCTTTTCTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M. Fatima Bonaldo. RNA was kindly provided by Dr. Minoru Ko (Wayne State University)."
 ORIGIN
 Query Match 45.3%; Score 27.2; DB 7; Length 451;
 Best Local Similarity 67.9%; Pred. No. 78;
 Matches 38; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
 Oy 5 GTACAGCGCCACCGCATGCGCTGACGAGCGCTCCGATACGAGCCGGCGCT 60
 |||||
 Db 108 GTACGCGCGGACCTTACGCTGCTTCACTGCGACGAAACAGAACCGCGCGCT 53
 |||||
 RESULT 15
 AZ933594 778 bp DNA linear GSS 24-APR-2001
 LOCUS BJ_Ba0001C21f B. japonicum BAC library Bradyrhizobium japonicum
 DEFINITION genomic, genomic survey sequence.
 ACCESSION AZ933594
 VERSION AZ933594
 KEYWORDS GSS.
 SOURCE Bradyrhizobium japonicum
 ORGANISM Bradyrhizobium japonicum
 Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Bradyrhizobiaceae; Bradyrhizobium.
 REFERENCE 1 (bases 1 to 778)

AUTHORS Tomkins, J.P., Wood, T.C., Stacey, M.G., Loh, J.T., Judd, A., Golcochea, J.L., Stacey, G., Sadowsky, M.J. and Wing, R.A.
 TITLE A marker-dense, sequence-ready map of the Bradyrhizobium japonicum genome
 JOURNAL Genome Res. 11 (8), 1434-1440 (2001)
 MEDLINE 21376150
 PUBMED 11483585
 COMMENT Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Class: BAC ends
 High quality sequence stop: 716.
 FEATURES
 source
 1..778
 /organism="Bradyrhizobium japonicum"
 /mol_type="genomic DNA"
 /strain="USDA110"
 /db_xref="taxon:375"
 /lab_host="E. coli"
 /clone_lib="B. japonicum BAC library"
 /note="Vector: pIndigo536; Site_1: HindIII"
 ORIGIN
 Query Match 45.3%; Score 27.2; DB 8; Length 778;
 Best Local Similarity 67.9%; Pred. No. 75;
 Matches 38; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
 Oy 2 GTATACAGCGCCACCGCATGCGCTCAGAGCGCTCCGATACGAGACCGCGG 57
 |||||
 Db 530 GTTGAAGAAGACTTCTGACCGCGCATCAGAGCTTTTCGGCTATGCGACCGCGG 585
 |||||

Search completed: January 16, 2005, 02:55:44
 Job time : 1554.6 secs

This Page Blank (uspto)

XX Sequence 60 BP; 12 A; 22 C; 19 G; 7 T; 0 U; 0 Other;
SQ

Query Match 100.0%; Score 60; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 3.6e-11;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGTAGTACAGCGCCACCCGACATGGCGTTCACGAGCGCTCCGGATACGAGACCGCGCCT 60
|||
DB 1 GGTAGTACAGCGCCACCCGACATGGCGTTCACGAGCGCTCCGGATACGAGACCGCGCCT 60
|||

RESULT 2

AAV48322

ID AAV48322 standard; DNA; 60 BP.

XX AAV48322;

DT 20-NOV-1998 (first entry)

DE Nucleotide sequence encoding MUC1 tandem repeat unit R7.

KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
tumour; tumour-associated antigen.

OS Homo sapiens.

PN WO9837095-A2.

PD 27-AUG-1998.

PF 24-FEB-1998; 98WO-US003693.

PR 24-FEB-1997; 97US-0038253P.

PA (THER-) THERION BIOLOGICS CORP.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (DAND) DANA FARBER CANCER INST INC.

PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;

DR WPI; 1998-467492/40.

PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS Disclosure; Page 11; 42pp; English.

CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
(RPV). The RPV was used in a pharmaceutical composition also containing
an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
does not undergo significant genetic deletion, thereby providing an
unexpectedly stable and immunogenic pox virus. They can be used to
prevent or treat tumours expressing MUC1 tumour-associated antigens

SQ Sequence 60 BP; 14 A; 24 C; 16 G; 6 T; 0 U; 0 Other;

Query Match 63.7%; Score 38.2; DB 2; Length 60;
Best Local Similarity 78.0%; Pred. No. 0.0011;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

OY 1 GGTAGTACAGCGCCACCCGACATGGCGTTCACGAGCGCTCCGGATACGAGACCGCGCCT 59
|||
DB 1 GGAAGTACCGCTCCACCTGCACACGGGGTCAACAAGCGCGCCAGACACTGACCTGCGCC 59
|||

RESULT 3

AAV48317

ID AAV48317 standard; DNA; 60 BP.

XX AAV48317;

XX 20-NOV-1998 (first entry)
DT
XX

DE Nucleotide sequence encoding MUC1 tandem repeat unit R2.

KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
tumour; tumour-associated antigen.

OS Homo sapiens.

PN WO9837095-A2.

PD 27-AUG-1998.

PF 24-FEB-1998; 98WO-US003693.

PR 24-FEB-1997; 97US-0038253P.

PA (THER-) THERION BIOLOGICS CORP.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (DAND) DANA FARBER CANCER INST INC.

PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;

DR WPI; 1998-467492/40.

PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS Disclosure; Page 11; 42pp; English.

CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
(RPV). The RPV was used in a pharmaceutical composition also containing
an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
does not undergo significant genetic deletion, thereby providing an
unexpectedly stable and immunogenic pox virus. They can be used to
prevent or treat tumours expressing MUC1 tumour-associated antigens

Query Match 62.7%; Score 37.6; DB 2; Length 60;
Best Local Similarity 76.7%; Pred. No. 0.0018;
Matches 46; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

OY 1 GGTAGTACAGCGCCACCCGACATGGCGTTCACGAGCGCTCCGGATACGAGACCGCGCCT 60
|||
DB 1 GGCAGTACTGCACCAACCGGACATGGCGTAAACATCAGCACCTGATACAAGACCTGACCT 60
|||

RESULT 4

AAV48321

ID AAV48321 standard; DNA; 60 BP.

XX AAV48321;

DT 20-NOV-1998 (first entry)

DE Nucleotide sequence encoding MUC1 tandem repeat unit R6.

KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
tumour; tumour-associated antigen.

OS Homo sapiens.

PN WO9837095-A2.

PD 27-AUG-1998.

PF 24-FEB-1998; 98WO-US003693.

PR 24-FEB-1997; 97US-0038253P.

```
XX (THER-) THERION BIOLOGICS CORP.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA (DAND ) DANA FARBER CANCER INST INC.
XX
PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX WPI; 1998-467492/40.
DR
XX New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS Disclosure; Page 11; 42pp; English.
XX
CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ Sequence 60 BP; 13 A; 27 C; 17 G; 3 T; 0 U; 0 Other;

Query Match      61.0%; Score 36.6; DB 2; Length 60;
Best Local Similarity 76.3%; Pred. No. 0.0039;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY      1 GGTAGTACAGCGCCACCGCATGGCGTCAAGAGCGCTCCGATACGAGACCGGCGCC 59
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      1 GGCAGCACCGCACCGCGCCGACACGCGGTCAAGCGCGCCAGACACTCGACCTGCGCC 59

RESULT 5
AAD00385
ID AAD00385 standard; DNA; 525 BP.
XX
AC AAD00385;
XX
DT 29-AUG-2000 (first entry)
XX
DE Human Mucin 1 (MUC-1) protein fragment encoding DNA #2.
XX
KW Human; Mucin 1; MUC-1; tumour; pMRS30 expression vector; anti-tumour;
KW therapy; immune response; cytostatic; vaccine; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..525
FT /*tag= a
FT /product= "MUC-1 protein fragment"
XX
PN WO200025827-A2.
XX
PD 11-MAY-2000.
XX
PF 18-OCT-1999; 99WO-EP007874.
XX
PR 30-OCT-1998; 98IT-MI002330.
XX
PA (MENA ) MENARINI RICERCHE SPA.
XX
PI Parente D, Di Massimo AM, De Santis R;
XX
DR WPI; 2000-365410/31.
DR P-PSDB; AAY71021.
XX
PT Composition containing one or more DNA molecules encoding fragments of a
PT Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
PT tumor therapy.
XX
```

```
PS Claim 16; Fig 2; 56pp; English.
XX
CC The present sequence is a DNA encoding a fragment of human Mucin 1 (MUC-
CC 1) antigenic protein which is overexpressed in tumour cells. The sequence
CC was obtained from BT20 tumour cells by reverse transcriptase-PCR and
CC corresponds to nucleotides 205-720 of the EMBL sequence J05581 with a
CC start codon and two stop codons. The present sequence is cloned into a
CC pMRS30 expression vector and used in pharmaceutical composition e.g.
CC vaccine for inducing an antigen-specific anti-tumour immune response.
CC Composition containing this DNA molecule is useful in anti-tumour therapy
CC of patients affected with tumours characterised by high MUC-1 expression
XX
SQ Sequence 525 BP; 111 A; 199 C; 128 G; 87 T; 0 U; 0 Other;

Query Match      61.0%; Score 36.6; DB 3; Length 525;
Best Local Similarity 76.3%; Pred. No. 0.0046;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY      1 GGTAGTACAGCGCCACCGCATGGCGTCAAGAGCGCTCCGATACGAGACCGGCGCC 59
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      256 GGAAGTACTGCTTCACACGACAGCAGCGGTGTTACTTCGGCTCCGATACGAGCGCGGCC 314

RESULT 6
AAD00391
ID AAD00391 standard; DNA; 891 BP.
XX
AC AAD00391;
XX
DT 15-SEP-2003 (revised)
DT 29-AUG-2000 (first entry)
XX
DE Ubiquitin-E. coli LacI-human Mucin 1 fusion protein encoding DNA #2.
XX
KW Ubiquitin; LacI; beta-galactosidase; fusion protein; human; Mucin 1;
KW MUC-1; tumour; pMRS30 expression vector; anti-tumour; therapy;
KW immune response; cytostatic; vaccine; ds.
XX
OS Homo sapiens.
OS Escherichia coli.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 1..891
FT /*tag= a
FT /product= "Ubiquitin-LacI-MUC-1 fusion protein"
XX
FT misc_feature 1..369
FT /*tag= b
FT /label= UBILacI_DNA
FT /note= "includes ubiquitin-E. coli lacI fusion DNA"
FT misc_feature 370..891
FT /*tag= c
FT /note= "Human MUC-1 partial DNA that corresponds to
FT nucleotides 205-720 of the EMBL sequence J05581 with two
FT stop codons"
XX
PN WO200025827-A2.
XX
PD 11-MAY-2000.
XX
PF 18-OCT-1999; 99WO-EP007874.
XX
PR 30-OCT-1998; 98IT-MI002330.
XX
PA (MENA ) MENARINI RICERCHE SPA.
XX
PI Parente D, Di Massimo AM, De Santis R;
XX
DR WPI; 2000-365410/31.
DR P-PSDB; AAY71027.
XX
PT Composition containing one or more DNA molecules encoding fragments of a
PT Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
```


PT tumor therapy.
XX
PS Claim 18; Fig 8; 56pp; English.
XX
CC The present sequence is a DNA encoding a fusion protein consisting of
CC human Mucin 1 (MUC-1) fragment fused to Ubilaci sequence at the N-
CC terminus. The Ubilaci sequence consists of ubiquitin from MCF7 cell line
CC and a portion of E. coli beta-galactosidase (Laci). MUC-1 is an antigenic
CC protein overexpressed in tumour cells. The present sequence is cloned
CC into a pMRS30 expression vector and used in pharmaceutical composition
CC e.g. vaccine for inducing an antigen-specific anti-tumour immune
CC response. Composition containing this DNA molecule is useful in anti-
CC tumour therapy of patients affected with tumours characterised by high
CC MUC-1 expression. (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ Sequence 891 BP; 203 A; 307 C; 225 G; 156 T; 0 U; 0 Other;

Query Match 61.0%; Score 36.6; DB 3; Length 891;
Best Local Similarity 76.3%; Pred. No. 0.0048;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

OY 1 GGTAAGTACAGCGCCACCGGACATGGCGTCACGAGCGCTCCGATACGAGACCGGCGCC 59
Db 622 GGAAGTACTGCTCCACCGACGACGCGTGTTCCTCGGCTCCGATACGAGCGCGCC 680

RESULT 7
AAD00388
ID AAD00388 standard; DNA; 1371 BP.
XX
AC AAD00388;
XX
DT 29-AUG-2000 (first entry)
XX
DE Human Mucin 1 (MUC-1) protein fragment encoding DNA #5.
XX
KW Human; Mucin 1; tumour; pMRS30 expression vector; anti-tumour;
KM therapy; immune response; cytostatic; vaccine; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1371
FT /*tag= a
FT /product= "MUC-1 protein fragment"
XX
PN WO200025827-A2.
XX
PD 11-MAY-2000.
XX
PF 18-OCT-1999; 99WO-EP007874.
XX
PR 30-OCT-1998; 98IT-MI002330.
XX
PA (MENA) MENARINI RICERCHE SPA.
XX
PI Parente D, Di Massimo AM, De Santis R;
XX
DR WPI; 2000-365410/31.
DR P-PSDB; AAY71024.
XX
PT Composition containing one or more DNA molecules encoding fragments of a
PT Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
PT tumor therapy.
XX
PS Claim 16; Fig 5; 56pp; English.
XX
CC The present sequence is a DNA encoding a fragment of human Mucin 1 (MUC-
CC 1) antigenic protein which is overexpressed in tumour cells. The sequence
CC was obtained by PCR from plasmids pMRS166, pMRS167, pMRS168 and pMRS169
CC which contain MUC-1 DNA from BT20 tumour cells. It corresponds to
CC nucleotides 136-1497 of the EMBL sequence J05581 with a start codon and
CC two stop codons. The present sequence is cloned into a pMRS30 expression

CC vector and used in pharmaceutical composition e.g. vaccine for inducing
CC an antigen-specific anti-tumour immune response. Composition containing
CC this DNA molecule is useful in anti-tumour therapy of patients affected
CC with tumours characterised by high MUC-1 expression
XX
SQ Sequence 1371 BP; 300 A; 450 C; 321 G; 300 T; 0 U; 0 Other;

Query Match 61.0%; Score 36.6; DB 3; Length 1371;
Best Local Similarity 76.3%; Pred. No. 0.0049;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

OY 1 GGTAAGTACAGCGCCACCGGACATGGCGTCACGAGCGCTCCGATACGAGACCGGCGCC 59
Db 325 GGAAGTACCGCTCCACCGACGACGCGTGTTCCTCGGCTCCGATACGAGCGCGCC 383

RESULT 8
AAD00394
ID AAD00394 standard; DNA; 1737 BP.
XX
AC AAD00394;
XX
DT 15-SEP-2003 (revised)
DT 29-AUG-2000 (first entry)
XX
DE Ubiquitin-E. coli Laci-human Mucin 1 fusion protein encoding DNA #5.
XX
KW Ubiquitin; Laci; beta-galactosidase; fusion protein; human; Mucin 1;
KM MUC-1; tumour; pMRS30 expression vector; anti-tumour; therapy;
XX
KW immune response; cytostatic; vaccine; ds.
XX
OS Homo sapiens.
OS Escherichia coli.
OS Chimeric.
XX

FH Key Location/Qualifiers
FT CDS 1..1737
FT /*tag= a
FT /product= "Ubiquitin-Laci-MUC-1 fusion protein"
FT misc_feature 1..369
FT /*tag= b
FT /label= Ubilaci DNA
FT /note= "Includes ubiquitin-E. coli Laci fusion DNA"
FT misc_feature 370..1737
FT /*tag= c
FT /note= "Human MUC-1 partial DNA that corresponds to
FT nucleotides 136-1497 of the EMBL sequence J05581 with two
FT stop codons"
XX
PN WO200025827-A2.
XX
PD 11-MAY-2000.
XX
PF 18-OCT-1999; 99WO-EP007874.
XX
PR 30-OCT-1998; 98IT-MI002330.
XX
PA (MENA) MENARINI RICERCHE SPA.
XX
PI Parente D, Di Massimo AM, De Santis R;
XX
DR WPI; 2000-365410/31.
DR P-PSDB; AAY71030.
XX
PT Composition containing one or more DNA molecules encoding fragments of a
PT Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
PT tumor therapy.
XX
PS Claim 18; Fig 11; 56pp; English.
XX
CC The present sequence is a DNA encoding a fusion protein consisting of
CC human Mucin 1 (MUC-1) fragment fused to Ubilaci sequence at the N-
CC terminus. The Ubilaci sequence consists of ubiquitin from MCF7 cell line

CC and a portion of E. coli beta-galactosidase (lacI). MUC-1 is an antigenic
CC protein overexpressed in tumour cells. The present sequence is cloned
CC into a pMRS30 expression vector and used in pharmaceutical composition
CC e.g. vaccine for inducing an antigen-specific anti-tumour immune
CC response. Composition containing this DNA molecule is useful in anti-
CC tumour therapy of patients affected with tumours characterised by high
CC MUC-1 expression. (Updated on 15-SEP-2003 to standardise OS field)

XX
SQ Sequence 1737 BP; 392 A; 558 C; 418 G; 369 T; 0 U; 0 Other;

Query Match

Best Local Similarity 61.0%; Score 36.6; DB 3; Length 1737;

Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

OY 1 GGTAAGTACAGCGCCACCGCACATGGCGTTCACGAGCGCTCCGATACGAGACCGGCGCC 59
Db 691 GGAAGTACCGCTCCACCGACACAGCGGTGTACCTCGGCTCCGATACGAGCGCGGCC 749

RESULT 9

AAV48320
ID AAV48320 standard; DNA; 60 BP.

XX
AC AAV48320;

DT 20-NOV-1998 (first entry)

XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit R5.

XX
KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KM tumour; tumour-associated antigen.

XX
OS Homo sapiens.

XX
PN WO9837095-A2.

XX
PD 27-AUG-1998.

XX
PF 24-FEB-1998; 98WO-US003693.

XX
PR 24-FEB-1997; 97US-0038253P.

XX
PA (THER-) THERION BIOLOGICS CORP.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (DAND) DANA FARBER CANCER INST INC.

PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;

XX
DR WPI; 1998-467492/40.

PT New recombinant pox virus for tumour therapy - comprises DNA encoding an

XX
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.

XX
PS Disclosure; Page 11; 42pp; English.

CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens

XX
SQ Sequence 60 BP; 9 A; 27 C; 15 G; 9 T; 0 U; 0 Other;

Query Match

Best Local Similarity 54.7%; Score 32.8; DB 2; Length 60;

Matches 43; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

OY 1 GGTAAGTACAGCGCCACCGCACATGGCGTTCACGAGCGCTCCGATACGAGACCGGCGCT 60
Db 1 GGTTGACGCGCCCGCTGCTCACGGGTAAACATCCGCCCGGATACCAAGCGGCCCT 60

RESULT 10

AAV48319
ID AAV48319 standard; DNA; 60 BP.

XX
AC AAV48319;

DT 20-NOV-1998 (first entry)

XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit R4.

XX
KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KM tumour; tumour-associated antigen.

XX
OS Homo sapiens.

XX
PN WO9837095-A2.

XX
PD 27-AUG-1998.

XX
PF 24-FEB-1998; 98WO-US003693.

XX
PR 24-FEB-1997; 97US-0038253P.

XX
PA (THER-) THERION BIOLOGICS CORP.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (DAND) DANA FARBER CANCER INST INC.

PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;

XX
DR WPI; 1998-467492/40.

PT New recombinant pox virus for tumour therapy - comprises DNA encoding an

XX
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.

XX
PS Disclosure; Page 11; 42pp; English.

CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens

XX
SQ Sequence 60 BP; 10 A; 23 C; 13 G; 14 T; 0 U; 0 Other;

Query Match 48.3%; Score 29; DB 2; Length 60;

Best Local Similarity 71.7%; Pred. No. 1.6;

Matches 38; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY 7 ACAGCGCCACCGCACATGGCGTTCACGAGCGCTCCGATACGAGACCGGCGCC 59
Db 7 ACAGCTCTCCCGCTCATGGGTTACTTCTGCTCCAGATACTCGCCAGCTCC 59

RESULT 11

ABK36806
ID ABK36806 standard; DNA; 72 BP.

XX
AC ABK36806;

DT 08-MAY-2002 (first entry)

XX
DE Human DNA encoding MUC1F segment 8.

XX
KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
KM viral infection; human immunodeficiency virus; melanoma;
KM bacterial infection; Salmonella; Legionella; parasitic infection;
KM Trypanosoma; Toxoplasma; Giardia; ds.

OS Homo sapiens.
XX
PN WO200190197-A1.
XX
PD 29-NOV-2001.
XX
PF 25-MAY-2001; 2001WO-AU000622.
XX
PR 26-MAY-2000; 2000AU-00007761.
XX
PA (AUSU) UNIV AUSTRALIAN NAT.
XX
PI Thomson SA, Ramshaw IA;
XX
DR WPI; 2002-147575/19.
DR P-PSDB; AAU84986.
XX
PT New synthetic polypeptides having several different segments of at least
PT one parent polypeptide linked together differently compared to the
PT linkage in the parent polypeptide, for inducing immune response against a
PT pathogen or cancer.
XX
PS Example 3; Fig 27; 364pp; English.
XX
CC The invention relates to a new synthetic polypeptide (I) comprising
CC several different segments of at least one parent polypeptide linked
CC together in a different relationship relative to their linkage in the
CC parent polypeptide to impede, abrogate or otherwise alter at least one
CC function associated with the parent polypeptide and for inducing an
CC immune response against a pathogen or cancer. Also included are a
CC synthetic polynucleotide encoding and a computer system for designing the
CC synthetic polypeptides. The synthetic polypeptides and polynucleotides
CC are referred to as a Savine. The synthetic polypeptide is useful for
CC modulating immune responses preferably directed against a pathogen or a
CC cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head
CC and neck, pancreas, prostate, stomach, bladder, kidney, bone liver,
CC oesophagus, brain, testicle, uterus), as potentiating agents.
CC Compositions comprising the polypeptide may be used in the treatment or
CC prophylaxis against viral (such as infections caused by HIV (human
CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,
CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic
CC Trypanosoma, Toxoplasma and Giardia) infections. The present sequence
CC encodes a peptide derived from a parent protein used to construct a
CC savine of the invention
XX
SQ Sequence 72 BP; 19 A; 29 C; 15 G; 9 T; 0 U; 0 Other;
XX
Query Match 48.0%; Score 28.8; DB 6; Length 72;
Best Local Similarity 69.6%; Pred. No. 1.9;
Matches 39; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
XX
QY 1 GGTA GTACAGCGCCACCCGCATGCGCTCAGCAGCGCTCCGGATACGAGACCGGC 56
Db 16 GGAAGCACACACCCCTCCCGCTCAGATGTGACAAAGCGCTCCCGATTAACAAGCGCG 71
XX
RESULT 12
ABK36828
ID ABK36828 standard; DNA; 16638 BP.
XX
AC ABK36828;
XX
DT 08-MAY-2002 (first entry)
XX
DE Human DNA for melanocyte differentiation antigens savine.
XX
KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
KW viral infection; human immunodeficiency virus; melanoma;
KW bacterial infection; Salmonella; Legionella; parasitic infection;
KW Trypanosoma; Toxoplasma; Giardia; ds.

XX Homo sapiens.
OS Synthetic.
XX
PN WO200190197-A1.
XX
PD 29-NOV-2001.
XX
PF 25-MAY-2001; 2001WO-AU000622.
XX
PR 26-MAY-2000; 2000AU-00007761.
XX
PA (AUSU) UNIV AUSTRALIAN NAT.
XX
PI Thomson SA, Ramshaw IA;
XX
DR WPI; 2002-147575/19.
DR P-PSDB; AAU85008.
XX
PT New synthetic polypeptides having several different segments of at least
PT one parent polypeptide linked together differently compared to the
PT linkage in the parent polypeptide, for inducing immune response against a
PT pathogen or cancer.
XX
PS Example 3; Fig 27; 364pp; English.
XX
CC The invention relates to a new synthetic polypeptide (I) comprising
CC several different segments of at least one parent polypeptide linked
CC together in a different relationship relative to their linkage in the
CC parent polypeptide to impede, abrogate or otherwise alter at least one
CC function associated with the parent polypeptide and for inducing an
CC immune response against a pathogen or cancer. Also included are a
CC synthetic polynucleotide encoding and a computer system for designing the
CC synthetic polypeptides. The synthetic polypeptides and polynucleotides
CC are referred to as a Savine. The synthetic polypeptide is useful for
CC modulating immune responses preferably directed against a pathogen or a
CC cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head
CC and neck, pancreas, prostate, stomach, bladder, kidney, bone liver,
CC oesophagus, brain, testicle, uterus), as potentiating agents.
CC Compositions comprising the polypeptide may be used in the treatment or
CC prophylaxis against viral (such as infections caused by HIV (human
CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,
CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic
CC Trypanosoma, Toxoplasma and Giardia) infections. The present sequence
CC encodes a savine protein of the invention
XX
SQ Sequence 16638 BP; 3840 A; 5297 C; 3944 G; 3557 T; 0 U; 0 Other;
XX
Query Match 48.0%; Score 28.8; DB 6; Length 16638;
Best Local Similarity 69.6%; Pred. No. 2.9;
Matches 39; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
XX
QY 1 GGTA GTACAGCGCCACCCGCATGCGCTCAGCAGCGCTCCGGATACGAGACCGGC 56
Db 4246 GGAAGCACACACCCCTCCCGCTCAGATGTGACAAAGCGCTCCCGATTAACAAGCGCG 4301
XX
RESULT 13
ADK68635/c
ID ADK68635 standard; cDNA; 156 BP.
XX
AC ADK68635;
XX
DT 06-MAY-2004 (first entry)
XX
DE HSP65-MUCL antigen CTL epitope related cDNA #3.
XX
KW Human breast cancer; Mycobacterium bovis; heat shock protein 65; BCG;
KW HSP65; cytotoxic T lymphocyte; CTL; MUCL; HSP65-MUCL antigen CTL epitope;
KW ss.

```
XX OS Unidentified.
XX PN CN1368384-A.
XX PD 11-SEP-2002.
XX PF 08-FEB-2001; 2001CN-00102614.
XX PR 08-FEB-2001; 2001CN-00102614.
XX PA (DIWEI-) DIWEIHUAYU BIO TECHNOLOGY CO LTD BEIJING.
XX PI Yu Y, Li H;
XX DR WPI; 2003-854662/80.
XX PT Genetically engineered vaccine of MUC-1 antigen for human breast cancer.
XX PS Example 2; Page 4 (Disclosure); 14pp; Chinese.
XX CC The invention relates to a method of preparation of a genetically
CC engineered vaccine for preventing and treating human breast cancer. The
CC method comprises fusing the coding gene of the Mycobacterium bovis heat
CC shock protein 65 (BCG HSP65) with the cytotoxic T lymphocyte (CTL)
CC epitope gene of MUC1 (antigen cell expressed by human breast cancer
CC cells) and the HSP65-MUC1 antigen CTL epitope is expressed in Escherichia
CC coli cells. This sequence represents DNA used in the method of the
CC invention.
XX SQ Sequence 156 BP; 37 A; 42 C; 58 G; 19 T; 0 U; 0 Other;
XX
XX Query Match 47.7%; Score 28.6; DB 10; Length 156;
XX Best Local Similarity 67.8%; Pred. No. 2.4;
XX Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 1 GGTAGTACAGCGCCACCGCACATGGCGCTCAGAGCGCTCCGATACGAGACCGCGCC 59
Db 132 GGTCTACCGCTCCGCCGCTCAGCGGTGTACCTTGCTCCGACACCCGTCGGCTCC 74
RESULT 14
ADK68629
ID ADK68629 standard; cDNA; 1800 BP.
XX AC ADK68629;
XX DT 06-MAY-2004 (first entry)
XX DE HSP65-MUC1 antigen CTL epitope related cDNA #1.
XX KW Human breast cancer; Mycobacterium bovis; heat shock protein 65; BCG;
KW HSP65; cytotoxic T lymphocyte; CTL; MUC1; HSP65-MUC1 antigen CTL epitope;
KW gene; ss.
XX OS Unidentified.
XX PN CN1368384-A.
XX PD 11-SEP-2002.
XX PF 08-FEB-2001; 2001CN-00102614.
XX PR 08-FEB-2001; 2001CN-00102614.
XX PA (DIWEI-) DIWEIHUAYU BIO TECHNOLOGY CO LTD BEIJING.
XX PI Yu Y, Li H;
XX DR WPI; 2003-854662/80.
XX DR P-PSDB; ADK68630.
XX PT Genetically engineered vaccine of MUC-1 antigen for human breast cancer.
```

```
XX PS Claim 1; SEQ ID NO 1; 14pp; Chinese.
XX CC The invention relates to a method of preparation of a genetically
CC engineered vaccine for preventing and treating human breast cancer. The
CC method comprises fusing the coding gene of the Mycobacterium bovis heat
CC shock protein 65 (BCG HSP65) with the cytotoxic T lymphocyte (CTL)
CC epitope gene of MUC1 (antigen cell expressed by human breast cancer
CC cells) and the HSP65-MUC1 antigen CTL epitope is expressed in Escherichia
CC coli cells. This sequence represents DNA used in the method of the
CC invention.
XX SQ Sequence 1800 BP; 347 A; 561 C; 616 G; 276 T; 0 U; 0 Other;
XX
XX Query Match 47.7%; Score 28.6; DB 10; Length 1800;
XX Best Local Similarity 67.8%; Pred. No. 2.8;
XX Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 1 GGTAGTACAGCGCCACCGCACATGGCGCTCAGAGCGCTCCGATACGAGACCGCGCC 59
Db 1639 GGTCTACCGCTCCGCCGCTCAGCGGTGTACCTTGCTCCGACACCCGTCGGCTCC 1697
RESULT 15
AD157666
ID AD157666 standard; cDNA; 4144 BP.
XX AC AD157666;
XX DT 22-APR-2004 (first entry)
XX DE Human breast specific nucleic acid (BSNA) #37.
XX KW Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
KW breast cancer; cytostatic.
XX OS Homo sapiens.
XX PN WO2003106648-A2.
XX PD 24-DEC-2003.
XX PF 16-JUN-2003; 2003WO-US018934.
XX PR 14-JUN-2002; 2002US-0389327P.
XX PA (DIAD-) DIADEXUS INC.
XX PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
XX DR WPI; 2004-082185/08.
XX DR P-PSDB; AD157740.
XX PT Novel isolated polypeptide comprising breast specific protein sequences,
XX PT useful for diagnosing or monitoring presence and metastases of breast
XX PT cancer in patient.
XX PS Claim 1; SEQ ID NO 37; 370pp; English.
XX CC The invention relates to human breast specific nucleic acids (BSNA) and
XX CC the breast specific proteins (BSP) they encode. The nucleic acids are
XX CC useful for determining the presence of a BSNA in a sample which involves
XX CC contacting the sample with a BSNA under conditions in which the BSNA will
XX CC selectively hybridise to a BSNA in the sample, and detecting the
XX CC hybridisation. The nucleic acids are useful for determining the presence
XX CC of a BSP in a sample which involves contacting the sample with suitable
XX CC reagent under conditions in which the reagent will selectively interact
XX CC with the BSP, and detecting the interaction of the reagent with a BSP in
XX CC the sample. The nucleic acids and proteins are useful for diagnosing or
XX CC monitoring the presence and metastases of breast cancer in a patient,
XX CC which involves determining an amount of nucleic acid or protein and
XX CC comparing the determined amount of nucleic acid or protein in the sample
XX CC of the patient to the amount of a breast specific marker in a normal
```

CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.

XX
SQ Sequence 4144 BP; 634 A; 1896 C; 1058 G; 552 T; 0 U; 4 Other;

Query Match 47.7%; Score 28.6; DB 12; Length 4144;
Best Local Similarity 69.8%; Pred. No. 3;
Matches 37; Conservative 1; Mismatches 15; Indels 0; Gaps 0;

QY 7 ACAGCGCCACCCCGACATGGCGTCAAGCGGCTCCGATACGAGACCGGCGCC 59
Db 699 ACCGCCCCCGCAGCCCAAGTGTCACTCGGCCCGACACCAAGCGCGCCC 751

Search completed: January 15, 2005, 20:36:12
Job time : 171.5 secs

Query Match 100.0%; Score 60; DB 6; Length 2297;
Best Local Similarity 100.0%; Pred. No. 6.1e-08;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTAGTACAGCGCCACCCGACATGGCGGTCTACGAGCGGCTCCGGATACGAGACCGCGCCT 60
|||||
Db 886 GGTAGTACAGCGCCACCCGACATGGCGGTCTACGAGCGGCTCCGGATACGAGACCGCGCCT 945
|||||

RESULT 2
BD225141
LOCUS
DEFINITION BD225141 525 bp DNA linear PAT 17-JUL-2003
Medicinal composition having antitumor effect and containing DNA
encoding antigenic protein.
BD225141
ACCESSION BD225141.1 GI:33034911
VERSION BD225141.1 GI:33034911
KEYWORDS JP 2002528519-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 525)
Pallente, D., Massimo, A.M.D. and Desantis, R.
Medicinal composition having antitumor effect and containing DNA
encoding antigenic protein
Patent: JP 2002528519-A 2 03-SEP-2002;
MENARINI RICERCHE SPA
OS Homo sapiens (human)
PN JP 2002528519-A/2
PD 03-SEP-2002
PF 18-OCT-1999 JP 2000579265
PR 30-OCT-1998 IT MI98A002330
PI DINO PALLENTE, ANNA MARIA D MASSIMO, RITA DESANTIS PC
A61K38/00, A61K35/76, A61K39/00, A61K48/00, A61P35/00, C12N15/09, PC
A61K37/02,
PC C12N15/00
CC Medicinal composition having antitumor effect and containing
CC DNA encoding
CC antigenic protein
FH Key Location/Qualifiers
FT source 1..525
Location/Qualifiers
1..525
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 61.0%; Score 36.6; DB 6; Length 525;
Best Local Similarity 76.3%; Pred. No. 0.74;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 GGTAGTACAGCGCCACCCGACATGGCGGTCTACGAGCGGCTCCGGATACGAGACCGCGC 59
|||||
Db 256 GGAAGTACTGCTCCACCGACACACGGGTGTACCTCGGCTCCGGATACCAAGCGCGCC 314
|||||

RESULT 3
BD225147
LOCUS
DEFINITION BD225147 891 bp DNA linear PAT 17-JUL-2003
Medicinal composition having antitumor effect and containing DNA
encoding antigenic protein.
BD225147
ACCESSION BD225147.1 GI:33034917
VERSION BD225147.1 GI:33034917
KEYWORDS JP 2002528519-A/8.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 891)
Pallente, D., Massimo, A.M.D. and Desantis, R.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
OS Homo sapiens (human)
PN JP 2002528519-A/5
PD 03-SEP-2002
PF 18-OCT-1999 JP 2000579265
PR 30-OCT-1998 IT MI98A002330
PI DINO PALLENTE, ANNA MARIA D MASSIMO, RITA DESANTIS PC
A61K38/00, A61K35/76, A61K39/00, A61K48/00, A61P35/00, C12N15/09, PC
A61K37/02,
PC C12N15/00
CC Medicinal composition having antitumor effect and containing
CC DNA encoding
CC antigenic protein
FH Key Location/Qualifiers
FT source 1..1371
Location/Qualifiers
1..1371
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source

TITLE
JOURNAL
COMMENT
OS Homo sapiens (human)
PN JP 2002528519-A 8 03-SEP-2002;
MENARINI RICERCHE SPA
OS Homo sapiens (human)
PN JP 2002528519-A/8
PD 03-SEP-2002
PF 18-OCT-1999 JP 2000579265
PR 30-OCT-1998 IT MI98A002330
PI DINO PALLENTE, ANNA MARIA D MASSIMO, RITA DESANTIS PC
A61K38/00, A61K35/76, A61K39/00, A61K48/00, A61P35/00, C12N15/09, PC
A61K37/02,
PC C12N15/00
CC Medicinal composition having antitumor effect and containing
CC DNA encoding
CC antigenic protein
FH Key Location/Qualifiers
FT source 1..891
Location/Qualifiers
1..891
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source

QY 1 GGTAGTACAGCGCCACCCGACATGGCGGTCTACGAGCGGCTCCGGATACGAGACCGCGCC 59
|||||
Db 622 GGAAGTACTGCTCCACCGACACACGGGTGTACCTCGGCTCCGGATACCAAGCGCGCC 680
|||||

RESULT 4
BD225144
LOCUS
DEFINITION BD225144 1371 bp DNA linear PAT 17-JUL-2003
Medicinal composition having antitumor effect and containing DNA
encoding antigenic protein.
BD225144
ACCESSION BD225144.1 GI:33034914
VERSION BD225144.1 GI:33034914
KEYWORDS JP 2002528519-A/5.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 1371)
Pallente, D., Massimo, A.M.D. and Desantis, R.
Medicinal composition having antitumor effect and containing DNA
encoding antigenic protein
Patent: JP 2002528519-A 5 03-SEP-2002;
MENARINI RICERCHE SPA
OS Homo sapiens (human)
PN JP 2002528519-A/5
PD 03-SEP-2002
PF 18-OCT-1999 JP 2000579265
PR 30-OCT-1998 IT MI98A002330
PI DINO PALLENTE, ANNA MARIA D MASSIMO, RITA DESANTIS PC
A61K38/00, A61K35/76, A61K39/00, A61K48/00, A61P35/00, C12N15/09, PC
A61K37/02,
PC C12N15/00
CC Medicinal composition having antitumor effect and containing
CC DNA encoding
CC antigenic protein
FH Key Location/Qualifiers
FT source 1..1371
Location/Qualifiers
1..1371
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source

ORIGIN

Query Match 61.0%; Score 36.6; DB 6; Length 1371;
Best Local Similarity 76.3%; Pred. No. 0.68;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 GGTAGTACAGCGCCACCCGCACATGGCGTCAAGAGCGCTCCGGATACGAGACCGCGCC 59
|||
Db 325 GGAAGTACCGCTCCACCACGACACACGGTGTACCTCGGCTCCGGATACGAGCGCGCC 383

RESULT 5

BD225150 1737 bp DNA linear PAT 17-JUL-2003
LOCUS
DEFINITION Medicinal composition having anticumor effect and containing DNA
encoding antigenic protein.
ACCESSION BD225150
VERSION BD225150.1 GI:33034920
KEYWORDS JP 2002528519-A/11.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
REFERENCE 1 (bases 1 to 1737)
AUTHORS Pallente,D., Massimo,A.M.D. and Desantis,R.
TITLE Medicinal composition having anticumor effect and containing DNA
encoding antigenic protein
Patent: JP 2002528519-A 11 03-SEP-2002;
JOURNAL MENARINI RICERCHE SPA
COMMENT OS Homo sapiens (human)
PN JP 2002528519-A/11
PD 03-SEP-2002 JP 2000579265
PF 18-OCT-1999 JP 2000579265
PR 30-OCT-1998 IT MI98A002330
PI DINO PALLEENTE,ANNA MARIA D MASSIMO,RITA DESANTIS PC
A61K38/00,A61K35/76,A61K39/00,A61K48/00,A61P35/00,C12N15/09, PC
A61K37/02,
PC C12N15/00
CC Medicinal composition having anticumor effect and containing
DNA encoding
CC antigenic protein
FH Key Location/Qualifiers
FT source 1..1737
FT location/Qualifiers
1..1737
/organism='Homo sapiens (human)'.
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'

ORIGIN

Query Match 61.0%; Score 36.6; DB 6; Length 1737;
Best Local Similarity 76.3%; Pred. No. 0.67;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 GGTAGTACAGCGCCACCCGCACATGGCGTCAAGAGCGCTCCGGATACGAGACCGCGCC 59
|||
Db 691 GGAAGTACCGCTCCACCACGACACACGGTGTACCTCGGCTCCGGATACGAGCGCGCC 749

RESULT 6

AX192396 120 bp DNA linear PAT 15-AUG-2001
LOCUS
DEFINITION Sequence 2 from Patent EP1103623.
ACCESSION AX192396
VERSION AX192396.1 GI:15210363
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
REFERENCE 1
AUTHORS Taylor-Papadimitriou,J., Burchell,J. and Gendler,S.

TITLE Human mucin core protein: nucleic acid probes, peptide fragments and antibodies thereto, and uses thereof in diagnostic and therapeutic methods

JOURNAL Patent: EP 1103623-A 2 30-MAY-2001;
IMPERIAL CANCER RESEARCH TECHNOLOGY LIMITED (GB)

FEATURES
source location/Qualifiers
1..120
/organism='Homo sapiens'
/mol_type='unassigned DNA'
/db_xref='taxon:9606'

ORIGIN

Query Match 45.7%; Score 27.4; DB 6; Length 120;
Best Local Similarity 69.8%; Pred. No. 4.9e+02;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACAGCGCCACCCGCACATGGCGTCAAGAGCGCTCCGGATACGAGACCGCGCC 59
|||
Db 81 ACCGCCCCCAGCCACGAGTGTACCTCGGCCCCGACACACGCGCGCCCC 29

RESULT 7

BD000571 120 bp DNA linear PAT 31-JAN-2002
LOCUS
DEFINITION Human polymorphic epithelial mucin core protein and nucleic acid
encoding the protein.
ACCESSION BD000571
VERSION BD000571.1 GI:18623684
KEYWORDS JP 2000333675-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
REFERENCE 1 (bases 1 to 120)
AUTHORS Papadimitrov,J.T., Jendora,S. and Bachieru,J.
TITLE Human polymorphic epithelial mucin core protein and nucleic acid
encoding the protein
Patent: JP 2000333675-A 2 05-DEC-2000;
JOURNAL IMPERIAL CANCER RESEARCH TECHNOLOGY LTD
COMMENT OS Homo sapiens (human)
PN JP 2000333675-A/2
PD 05-DEC-2000
PF 26-APR-2000 JP 2000125724
PR 07-JAN-1987 GB 8700269,07-JAN-1987 GB 8700279 PR
22-APR-1987 US 041306,09-NOV-1987 GB 8726172 PI JOYCE
TAYLOR PAPADIMITROV,SANDRA JENDORA,JOY BACHIERU PC
C12N15/02,A61K38/00,A61K39/395,A61K39/395,A61K49/00,A61P35/00, PC
C07K14/47,
PC C07K16/44,C12N5/10,C12P21/08/(C12N15/02,C12R1:91),(C12N5/10,
PC C12R1:91),
PC C12N15/00,A61K37/02,C12N5/00,(C12N15/00,C12R1:91),(C12N5/00,
PC C12R1:91)
CC
FH Key Location/Qualifiers
FT source 1..120
FT location/Qualifiers
1..120
/organism='Homo sapiens (human)'.
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'

ORIGIN

Query Match 45.7%; Score 27.4; DB 6; Length 120;
Best Local Similarity 69.8%; Pred. No. 4.9e+02;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACAGCGCCACCCGCACATGGCGTCAAGAGCGCTCCGGATACGAGACCGCGCC 59
|||
Db 81 ACCGCCCCCAGCCACGAGTGTACCTCGGCCCCGACACACGCGCGCCCC 29

RESULT 8

E08763
LOCUS E08763 180 bp RNA linear PAT 29-SEP-1997
DEFINITION cDNA encoding N-terminal fragment of human glycoprotein 39.
ACCESSION E08763
VERSION E08763.1 GI:2176875
KEYWORDS JP 1995051065-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 180)
Masuzawa,Y., Muramatsu,T. and Miyauchi,T.
GLYCOPROTEIN 39 GENE
Patent: JP 1995051065-A 1 28-FEB-1995;
NIPPON KOUTAI KENKYUSHO:KK, KAGOSHIMA UNIV
OS Homo sapiens (human)
PN JP 1995051065-A/1
PD 28-FEB-1995
PF 21-FEB-1992 JP 1992035085
PI MASUZAWA YASUSHI, MURAMATSU TAKASHI, MIYAUCHI TERUO PC
C12N15/09,G01N33/50//C07K14/00,C12P21/02;
CC strandedness: Double;
CC topology: linear;
CC hypothetical: No;
CC anti-sense: No;
CC Feature is identified by similarity;
FH Key Location/Qualifiers
FT source 1..180
FT /organism='Homo sapiens'
FT /cell_type='gastric cancer'
FT /cell_line='KATO-3'
FT /clone_lib='lambda gtl1 KATO-3 cDNA library'
FT /product='N-terminal fragment of human FT
FT glycoprotein 39',
FT repeat_region 1..180
FT repeat unit 1..60.
FT Location/Qualifiers
1..180
/organism="Homo sapiens"
/mol_type="genomic RNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 45.7%; Score 27.4; DB 6; Length 180;
Best Local Similarity 69.8%; Pred. No. 4.7e+02;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 7 ACAGCGCCACCCGACATGCGGTACGACGCGCTCCGATACGAGACCGGCGCC 59
Db 127 ACCGCGCCCGACGCCACGCGTGTCACTCGGCCCCCGACACACGCGCGGCC 179
RESULT 9
AF423031 1414 bp mRNA linear SYN 10-JUL-2003
LOCUS AF423031
DEFINITION Synthetic construct Homo sapiens mucin variant MUC1-CT58 (MUC1)
ACCESSION AF423031
VERSION AF423031.1 GI:19338621
KEYWORDS
SOURCE
ORGANISM
synthetic construct
synthetic construct
artificial sequences.
1 (bases 1 to 1414)
Hinojosa-Kurtzberg,A.M., Johansson,M.E., Madsen,C.S., Hansson,G.C.
and Gendler,S.J.
Novel MUC1 splice variants contribute to mucin overexpression in
CFTR-deficient mice
Am. J. Physiol. Gastrointest. Liver Physiol. 284 (5), G853-G862
REFERENCE
AUTHORS
TITLE
JOURNAL

MEDLINE (2003)
22570517
PUBMED 12529261
REFERENCE 2 (bases 1 to 1414)
AUTHORS Hinojosa-Kurtzberg,A.M. and Gendler,S.J.
TITLE Direct Submission
JOURNAL Submitted (21-SEP-2001) Biochemistry, Mayo Clinic, 13400 E. Shea
Boulevard, Scottsdale, AZ 85259, USA
FEATURES
source location/Qualifiers
1..1414
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/note="Homo sapiens gene in transgenic Mus musculus
C57BL/6; isolated from intestinal mucosa"
1..1414
/gene="MUC1"
1..1386
/gene="MUC1"
/note="alternatively spliced; contains exon 6b resulting
in variant carboxy-terminal domain; lacks sites for
beta-catenin and Grb2 interactions; derived from Homo
sapiens"
/codon_start=1
/transl_table=11
/product="mucin variant MUC1-CT58"
/protein_id="AAL86735.1"
/db_xref="GI:19338622"
/translation="MTPGTQSPFFLLTLTLTVTVTSGHASTPGREKETSATORSS
VPSSTEKNVSMSTSVLSHSPGSGSSTTQGDVTLADATBPASGSAATWGQDVTSVP
VTRPALGSTPPAHADVTSAPDNKPAPGSTAPPAHGVTSAPDTRAPAGSTAPPAHGVTS
APDNRPALGSTAPPAHGVTSAPDNKPAPGSTAPPAHGVTSAPDTRAPAGSTAPPAHGVTS
SDPTTLASHSTKTDASTHSTVPPLTSSNHSSTPSTGVSFFLSFHSNLSGNS
SLEDPTDYYOELORDISEMFLQIKOGFGLSNIKRPGSVVQVLTIAFREGLTV
HDVETQFNQYKTBASRYNLTISDVSVDPFPFSAQAGAGVPGWIALVLVCLVA
LAIVYLIALAVQCQRKNYGLDIFPARDTYHPMSERYPTHGRYVPPSSTDRSPE
KEWRVDRERLA"
ORIGIN
Query Match 45.7%; Score 27.4; DB 12; Length 1414;
Best Local Similarity 69.8%; Pred. No. 4e+02;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 7 ACAGCGCCACCCGACATGCGGTACGACGCGCTCCGATACGAGACCGGCGCC 59
Db 391 ACCGCGCCCGACGCCACGCGTGTCACTCGGCCCCCGACACACGCGCGGCC 443
RESULT 10
CQ715242 1455 bp DNA linear PAT 03-FEB-2004
LOCUS CQ715242
DEFINITION Sequence 1176 from Patent WO02068579.
ACCESSION CQ715242
VERSION CQ715242.1 GI:42276099
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1
Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
Patent: WO 02068579-A 1176 06-SEP-2002;
PE Corporation (NY) (US)
location/Qualifiers
1..1455
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
FEATURES
source
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
ORIGIN

Query Match	45.7%;	Score 27.4;	DB 6;	Length 1455;
Best Local Similarity	69.8%;	Pred. No. 4e+02;		
Matches 37;	Conservative 0;	Mismatches 16;	Indels 0;	Gaps 0;
Qy	7	ACACGGCCACCCGCATGCGCTTCACGAGCGCTCCGGATACGAGACCGGGCC	59	
Db	418	ACCGCCCCCCACGCGGTTCACCTCGGCCCCGACACCGCGGGCCCC	470	
RESULT 11				
AX959914				
LOCUS	AX959914	1457 bp	DNA	linear
DEFINITION	Sequence 19 from Patent WO03100060.			PAT 14-JAN-2004
ACCESSION	AX959914			
VERSION	AX959914.1	GI:40880143		
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE				
AUTHORS	Burden,N.G., Ellis,J.H. and Hamblin,P.A.			
TITLE	Muc-1 antigen with reduced number of vntx repeat units			
JOURNAL	Patent: WO 03100060-A 19 04-DEC-2003;			
	GLAXO GROUP LIMITED (GB)			
FEATURES				
source	1. .1457			
	/organism="Homo sapiens"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:9606"			
ORIGIN				
Query Match	45.7%;	Score 27.4;	DB 6;	Length 1457;
Best Local Similarity	69.8%;	Pred. No. 4e+02;		
Matches 37;	Conservative 0;	Mismatches 16;	Indels 0;	Gaps 0;
Qy	7	ACACGGCCACCCGCATGCGCTTCACGAGCGCTCCGGATACGAGACCGGGCC	59	
Db	409	ACCGCCCCCCACGCGGTTCACCTCGGCCCCGACACCGCGGGCCCC	461	
RESULT 12				
AX093798				
LOCUS	AX093798	1572 bp	DNA	linear
DEFINITION	Sequence 1 from Patent WO0118035.			PAT 30-MAR-2001
ACCESSION	AX093798			
VERSION	AX093798.1	GI:13510057		
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE				
AUTHORS	Taylor-Papadimitriou,J., Heukamp,L.C., Offringa,R., Melief,C.J.,			
TITLE	Acres,B. and Thomas,M.			
JOURNAL	Muc-1 derived peptides			
	Patent: WO 0118035-A 1 15-MAR-2001;			
	TRANSGENE S.A. (FR) ; IMPERIAL CANCER RESEARCH TECHNOLOGY LIMITED			
	(GB)			
FEATURES				
source	Location/Qualifiers			
	1. .1572			
	/organism="Homo sapiens"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:9606"			
	58. .1545			
	/note="unnamed protein product"			
	/codon_start=1			
	/protein_id="CAC35422.1"			
	/db_xref="GI:13510058"			
	/translation="MTPTGTSPPFLLLLTLVTVVTSGLHSSRPGEKETSATQRSS			
	VPSSTEKNVAVMTSSVLSHSPGSGSSTOGQDVTIAPATEPASGSAATWGQDVTSP			
	VTPALGSTTPPAHDVTSAPDNKPAGGSTAPPAHGVTSAPDTRPBGSTAPPAHGVTS			
	APDTRPAGSTAPPAHGVTSAPDNKPALGSTAPVHNVTSASGSASGASTLVHNGTS			

ORIGIN	ARATTPPASKSPSPISPHSHDPTPTLASHSTKTDASSTHSTVPLTSSNHSSTSPQ STGVSEFFLSFPISNLOFNSSLSDPSTDYVELORDISEMFLQIKOGGFLGLSNIKF RPGSVVQLTLAFREGTINVADVETQFNQKTEAASRYNLITSDVSVHVFPFSAQS GAGVGWGIALLVLVCVLAIVLYLIALAVCOCRRKNYGOLDIFPARDTYHPMSEYP TYHTHGRTVPSSSTDRSPYEKVASAGNGSSSLSTYNPAVAATSANL"				
Oy	7 ACACGGCCACC	GACATGCCTCACGACGCGCTCCGGATACGAGACCGGCC	59		
Dbl	508 ACCGCCCCCGCCCGCCACGCGTGTCACCTCGGCCCGGACACCAAGCGGCC	560			
<hr/>					
RESULT 13					
LOCUS	CQ771290	1721 bp	DNA	linear	PAT 04-MAR-2004
DEFINITION	Sequence 1 from Patent WO2004009632.				
ACCESSION	CQ771290				
VERSION	CQ771290.1	GI:45125397			
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 Goletz,S. and Karsten,U. Method for the production of an immunostimulating mucin (muc1) Patent: WO 2004009632-A 1 29-JAN-2004;				
AUTHORS	Nemod Immuntherapie AG (DE)				
TITLE	Location/Qualifiers				
JOURNAL	1. .1721 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"				
FEATURES	source				
<hr/>					
ORIGIN					
Query Match	45.7%;	Score 27.4;	DB 6;	Length 1572;	
Best Local Similarity	69.8%;	Pred. No. 3.9e+02;			
Matches	37;	Conservative	0;	Mismatches	16; Indels 0; Gaps 0;
Oy	7 ACACGGCCACC	GACATGCCTCACGACGCGCTCCGGATACGAGACCGGCC	59		
Dbl	508 ACCGCCCCCGCCCGCCACGCGTGTCACCTCGGCCCGGACACCAAGCGGCC	560			
<hr/>					
RESULT 14					
LOCUS	AR492306	1721 bp	DNA	linear	PAT 15-MAY-2004
DEFINITION	Sequence 3 from patent US 6716627.				
ACCESSION	AR492306				
VERSION	AR492306.1	GI:47260880			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 1721)				
AUTHORS	Dobie,K.W.				
TITLE	Antisense modulation of mucin 1, transmembrane expression				
JOURNAL	Patent: US 6716627-A 3 06-APR-2004;				
FEATURES	Location/Qualifiers				
source	1. .1721 /organism="unknown" /mol_type="genomic DNA"				
<hr/>					
ORIGIN					
Query Match	45.7%;	Score 27.4;	DB 6;	Length 1721;	
Best Local Similarity	69.8%;	Pred. No. 3.9e+02;			
Matches	37;	Conservative	0;	Mismatches	16; Indels 0; Gaps 0;
Oy	7 ACACGGCCACC	GACATGCCTCACGACGCGCTCCGGATACGAGACCGGCC	59		

Db 568 ACCCGCGCCCGCAGCCCGCAGGTGTCACTCGGCCCCGAGACACGAGCGCGCCCC 620

RESULT 15
AX335860

LOCUS AX335860 1721 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 6369 from Patent WO0194629.
ACCESSION AX335860
VERSION AX335860.1 GI:18126579
KEYWORDS

SOURCE
ORGANISM Homo sapiens (human)

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

TITLE
1 Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrigan, S., Soppet, D.R. and Weaver, Z.
Cancer gene determination and therapeutic screening using signature

JOURNAL
Patent: WO 0194629-A 6369 13-DEC-2001,
Avalon Pharmaceuticals (US)

FEATURES
source location/Qualifiers
1..1721

ORIGIN
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Query Match 45.7%; Score 27.4; DB 6; Length 1721;
Best Local Similarity 69.8%; Pred. No. 3.9e+02;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACAGCGCCACCCGACATGCGCTCAGAGCGCTCCGATACGAGACCGCGCC 59
568 ACCCGCGCCCGCAGCCCGCAGGTGTCACTCGGCCCCGAGACACGAGCGCGCCCC 620

Search completed: January 15, 2005, 22:36:55
Job time : 723.8 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 22:37:00 : Search time 183.4 Seconds
(without alignments)
1879.789 Million cell updates/sec

Title: US-10-057-136-11

Perfect score: 60

Sequence: 1 ggtcaacggcaccctccagc.....ccgacacccgtccagtcgcg 60

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	100.0	60	14 US-10-057-136-11	Sequence 11, Appl
2	60	100.0	2297	16 US-10-406-317-41	Sequence 41, Appl
3	38.2	63.7	60	14 US-10-057-136-5	Sequence 5, Appl
4	36.6	61.0	60	14 US-10-057-136-6	Sequence 6, Appl
5	36	60.0	60	14 US-10-057-136-2	Sequence 2, Appl
6	36	60.0	78	14 US-10-057-136-13	Sequence 13, Appl
7	36	60.0	1424	16 US-10-447-839A-75	Sequence 75, Appl
8	36	60.0	1428	16 US-10-447-839A-20	Sequence 20, Appl
9	36	60.0	1527	14 US-10-057-136-19	Sequence 19, Appl
10	36	60.0	1799	16 US-10-447-839A-19	Sequence 19, Appl
11	36	60.0	1804	9 US-09-964-824A-573	Sequence 573, App
12	36	60.0	1804	15 US-10-029-517-17	Sequence 17, Appl

13	36	60.0	1804	17	US-10-717-597-30	Sequence 30, Appl
14	36	60.0	1804	17	US-10-775-920-84	Sequence 84, Appl
15	36	60.0	1823	15	US-10-101-510-339	Sequence 339, App
16	36	60.0	4139	9	US-09-964-824A-105	Sequence 105, App
17	36	60.0	4139	9	US-09-964-824A-578	Sequence 578, App
18	36	60.0	4139	9	US-09-864-864-334	Sequence 334, App
19	36	60.0	4139	9	US-09-880-107-2121	Sequence 2121, App
20	36	60.0	4139	11	US-09-968-007A-751	Sequence 751, App
21	36	60.0	4139	14	US-10-171-311-157	Sequence 157, App
22	36	60.0	4139	15	US-10-177-293-310	Sequence 310, App
23	36	60.0	4139	16	US-10-440-464-155	Sequence 155, App
24	36	60.0	4139	17	US-10-734-564-53	Sequence 53, App
25	36	60.0	4139	17	US-10-775-920-80	Sequence 80, Appl
26	36	60.0	4139	17	US-10-775-920-85	Sequence 85, Appl
27	36	60.0	8181	16	US-10-447-839A-18	Sequence 18, Appl
28	36	60.0	8186	15	US-10-029-517-19	Sequence 19, Appl
29	36	60.0	8186	15	US-10-029-517-18	Sequence 18, Appl
30	34.4	57.3	572	17	US-10-775-920-93	Sequence 93, Appl
31	34.4	57.3	572	17	US-10-775-920-93	Sequence 280, App
32	34.4	57.3	1721	9	US-09-864-864-280	Sequence 224, App
33	34.4	57.3	1721	9	US-09-967-768A-224	Sequence 21, Appl
34	34.4	57.3	1721	14	US-10-247-703-21	Sequence 211, App
35	34.4	57.3	1721	14	US-10-097-340-211	Sequence 155, App
36	34.4	57.3	1721	14	US-10-171-311-155	Sequence 58, Appl
37	34.4	57.3	1721	15	US-10-007-926A-58	Sequence 3, Appl
38	34.4	57.3	1721	15	US-10-029-517-3	Sequence 775, App
39	34.4	57.3	1721	15	US-10-172-118-775	Sequence 775, App
40	34.4	57.3	1721	16	US-10-342-887-775	Sequence 88, Appl
41	34.4	57.3	1721	17	US-10-775-920-88	Sequence 12589, A
42	34.4	57.3	2026	14	US-10-198-846-12589	Sequence 87, Appl
43	34.4	57.3	2238	17	US-10-775-920-87	Sequence 103, Appl
44	34.4	57.3	2678	15	US-10-252-157-103	Sequence 38, Appl
45	33.6	56.0	518	14	US-10-247-703-38	

ALIGNMENTS

RESULT 1
US-10-057-136-11
: Sequence 11, Application US/10057136
: Publication No. US20030021770A1
: GENERAL INFORMATION:
: APPLICANT: SCHLON, JEFFREY
: APPLICANT: KANTOR, JUDITH
: APPLICANT: KUFE, DONALD
: APPLICANT: PANICALI, DENNIS
: TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
: FILE REFERENCE: 700953/47113C
: CURRENT APPLICATION NUMBER: US/10/057,136
: PRIOR FILING DATE: 2002-01-25
: PRIOR APPLICATION NUMBER: 09/366,670
: PRIOR FILING DATE: 1999-08-03
: PRIOR APPLICATION NUMBER: PCT/US98/03693
: PRIOR FILING DATE: 1998-02-24
: PRIOR APPLICATION NUMBER: 60/038,253
: NUMBER OF SEQ ID NOS: 20
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 11
: LENGTH: 60
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-057-136-11

Query Match 100.0%; Score 60; DB 14; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.2e-11;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGTCAACGCACCTCCAGACACGAGTCACGTTCACCCGACACCCGTCAGTCGCG 60
|||||

Db 1 GGTTCACGGCAGCCTCCAGCACACGAGAGTCAAGTCTGCACCCGACACCCGCTCCAGCTCCG 60

RESULT 2

US-10-406-317-41
; Sequence 41, Application US/10406317
; Publication No. US20040019195A1
; GENERAL INFORMATION:
; APPLICANT: Schlom, Jeffrey;
; APPLICANT: Hodge, James;
; APPLICANT: Panicali, Dennis
; TITLE OF INVENTION: A recombinant vector expressing multiple constitutulatory
; FILE REFERENCE: 38163-0189
; CURRENT APPLICATION NUMBER: US/10/406,317
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US/09/856,988
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: PCT/US99/26866
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/111,582
; PRIOR FILING DATE: 1998-12-09
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: VECTOR
US-10-406-317-41

Query Match 100.0%; Score 60; DB 16; Length 2297;
Best Local Similarity 100.0%; Pred. No. 1.3e-11;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCCTCCAGCACACGAGTCAAGTCTGCACCCGACACCCGCTCCAGCTCCG 60
Db 826 GGTTCACGGCAGCCTCCAGCACACGAGTCAAGTCTGCACCCGACACCCGCTCCAGCTCCG 885

RESULT 3

US-10-057-136-5
; Sequence 5, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-5

Query Match 63.7%; Score 38.2; DB 14; Length 60;
Best Local Similarity 78.0%; Pred. No. 0.00059;

Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCCTCCAGCACACGAGTCAAGTCTGCACCCGACACCCGCTCCAGCTCC 59
Db 1 GGATCCACCGCGCCGCTCCGACGAGTGAAGTCTGCGCGCCCGACACGCGCCCGCTCC 59

RESULT 4

US-10-057-136-6
; Sequence 6, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-6

Query Match 61.0%; Score 36.6; DB 14; Length 60;
Best Local Similarity 76.3%; Pred. No. 0.0021;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCCTCCAGCACACGAGTCAAGTCTGCACCCGACACCCGCTCCAGCTCC 59
Db 1 GGGTCAACAGCTCCTCCCGCTCATGGGGTACTTCTGCTCCAGATACTCGCCAGCTCC 59

RESULT 5

US-10-057-136-2
; Sequence 2, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-2

Query Match	60.0%	Score 36;	DB 14;	Length 60;
Best Local Similarity	75.0%	Pred. No. 0.0035;		
Matches 45; Conservative	0;	Mismatches 15;	Indels 0;	Gaps 0;

ORGANISM: Homo sapiens
US-10-057-136-19

Query Match
Best Local Similarity 60.0%; Score 36; DB 14; Length 1527;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTCAACGGCACCCTCCAGACACGAGTCACTGACCCGACACCCGTCAGTCCG 60
Db 226 GGCTCCACCGCCCCCGACGCGGTGTCACTCGGCCCGGACACGAGCGGCCCG 285

RESULT 10

US-10-447-839A-19
; Sequence 19, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald W.
; APPLICANT: Khatbada, Surrender
; APPLICANT: Weitman, Steven D.
; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT APPLICATION NUMBER: US/10/447,839A
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 1799
; TYPE: RNA
; ORGANISM: RNA
US-10-447-839A-19

Query Match
Best Local Similarity 60.0%; Score 36; DB 16; Length 1799;
Matches 42; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTCAACGGCACCCTCCAGACACGAGTCACTGACCCGACACCCGTCAGTCCG 60
Db 453 GGCUCACCGCCCCCGACGCGGTGTCACTCGGCCCGGACACGAGCGGCCCG 512

RESULT 11

US-09-964-824A-573
; Sequence 573, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horrigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964,824A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 573
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-824A-573

Query Match 60.0%; Score 36; DB 9; Length 1804;

Best Local Similarity 75.0%; Pred. No. 0.0036;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTCAACGGCACCCTCCAGACACGAGTCACTGACCCGACACCCGTCAGTCCG 60
Db 457 GGCTCCACCGCCCCCGACGCGGTGTCACTCGGCCCGGACACGAGCGGCCCG 516

RESULT 12

US-10-029-517-17
; Sequence 17, Application US/10029517
; Publication No. US20030148969A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 17
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (73)...(1500)
US-10-029-517-17

Query Match
Best Local Similarity 60.0%; Score 36; DB 15; Length 1804;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTCAACGGCACCCTCCAGACACGAGTCACTGACCCGACACCCGTCAGTCCG 60
Db 457 GGCTCCACCGCCCCCGACGCGGTGTCACTCGGCCCGGACACGAGCGGCCCG 516

RESULT 13

US-10-717-597-30
; Sequence 30, Application US/10717597
; Publication No. US20040110221A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael E.
; APPLICANT: Twine, Natalie C.
; APPLICANT: Doerner, Andrew J.
; APPLICANT: Trepicchio, William L.
; APPLICANT: Slonim, Donna K.
; APPLICANT: Stover, Jennifer A.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING RCC AND OTHER SOLID TUMORS
; FILE REFERENCE: AM101080L
; CURRENT APPLICATION NUMBER: US/10/717,597
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US 60/459,782
; PRIOR FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: US 60/427,982
; PRIOR FILING DATE: 2002-11-21
; NUMBER OF SEQ ID NOS: 4904
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-717-597-30

Query Match
Best Local Similarity 60.0%; Score 36; DB 17; Length 1804;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTCAACGGCACCCTCCAGACACGAGTCACTGACCCGACACCCGTCAGTCCG 60
Db 457 GGCTCCACCGCCCCCGACGCGGTGTCACTCGGCCCGGACACGAGCGGCCCG 516

RESULT 14

US-10-775-920-84
; Sequence 84, Application US/10775920
; Publication No. US20040175744A1
; GENERAL INFORMATION:
; APPLICANT: Mergen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; TITLE OF INVENTION: IN CERTAIN CANCERS
; FILE REFERENCE: Mergen - 0010B
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 84
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-84

Query Match 60.0%; Score 36; DB 17; Length 1804;
Best Local Similarity 75.0%; Pred. No. 0.0036;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTCAACGGCAGCTCCAGACACGAGTCACTGACCCGACACCCGTCAGTCCG 60
DB 457 GGCTCCACCGCCCCCGCCAGCCAGCGGTGTCACTCGCCCCGACACACAGCGCGCCCG 516

RESULT 15

US-10-101-510-339
; Sequence 339, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 339
; LENGTH: 1823
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-339

Query Match 60.0%; Score 36; DB 15; Length 1823;
Best Local Similarity 75.0%; Pred. No. 0.0036;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTCAACGGCAGCTCCAGACACGAGTCACTGACCCGACACCCGTCAGTCCG 60
DB 457 GGCTCCACCGCCCCCGCCAGCCAGCGGTGTCACTCGCCCCGACACACAGCGCGCCCG 516

Search completed: January 16, 2005, 09:30:31
Job time : 184.4 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: January 15, 2005, 19:50:13 : Search time 35.2 Seconds
(without alignments)
1211.572 Million cell updates/sec

Title: US-10-057-136-11

Perfect score: 60
Sequence: 1 ggttcacggcaccctccagc.....ccgacacccgctccagctccg 60

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	60.0	1804	US-10-029-517-17	Sequence 17, Appl
2	36	60.0	8186	US-10-029-517-19	Sequence 19, Appl
3	34.4	57.3	572	US-10-029-517-18	Sequence 18, Appl
4	34.4	57.3	1721	US-10-029-517-3	Sequence 3, Appl
5	33.6	56.0	518	US-10-029-517-101	Sequence 101, App
6	32.2	53.7	60	US-09-475-947A-246	Sequence 246, App
7	31.2	52.0	3343	US-10-029-517-102	Sequence 102, App
8	31.2	52.0	6192	US-08-479-537A-1	Sequence 1, Appl
9	31.2	52.0	6192	US-09-083-116-1	Sequence 1, Appl
10	31.2	52.0	6192	US-09-134-916A-1	Sequence 1, Appl
11	31.2	52.0	6449	US-08-479-537A-4	Sequence 4, Appl
12	31.2	52.0	6449	US-09-083-116-4	Sequence 4, Appl
13	31.2	52.0	6449	US-09-134-916A-4	Sequence 4, Appl
14	30.4	50.7	981	US-10-029-517-16	Sequence 16, Appl
15	28.6	47.7	519	US-09-646-028-42	Sequence 42, Appl
16	28.6	47.7	534	US-09-446-028-46	Sequence 46, Appl
17	25.4	42.3	6617	US-09-976-594-268	Sequence 268, App
18	24.6	41.0	2338	US-08-425-069-1	Sequence 1, Appl
19	24.6	41.0	2338	US-08-317-844B-1	Sequence 1, Appl
20	23.6	39.3	806	US-09-154-083-7	Sequence 7, Appl
21	23.6	39.3	2126	US-08-789-354-1	Sequence 1, Appl
22	23.6	39.3	2126	US-09-110-937-1	Sequence 1, Appl
23	23.6	39.3	2126	US-09-058-725B-1	Sequence 1, Appl
24	23.6	39.3	2126	US-09-232-857-1	Sequence 1, Appl
25	23.6	39.3	4403765	US-09-103-840A-2	Sequence 2, Appl
26	23.6	39.3	441529	US-09-103-840A-1	Sequence 1, Appl
27	23.4	39.0	534	US-09-621-976-15293	Sequence 15293, A

C	28	23.4	39.0	1894	3	US-09-329-350-32	Sequence 32, Appl
C	29	23.4	39.0	1894	4	US-08-841-636A-32	Sequence 32, Appl
C	30	23.2	38.7	2755	4	US-09-907-794A-90	Sequence 90, Appl
C	31	23.2	38.7	2755	4	US-09-905-125A-90	Sequence 90, Appl
C	32	23.2	38.7	2755	4	US-09-902-775A-90	Sequence 90, Appl
C	33	23.2	38.7	2755	4	US-09-906-700-90	Sequence 90, Appl
C	34	23.2	38.7	2755	4	US-10-140-002-353	Sequence 353, Appl
C	35	23.2	38.7	2755	4	US-09-903-603A-90	Sequence 90, Appl
C	36	23	38.3	7515	4	US-09-328-352-1695	Sequence 1695, Ap
C	37	22.8	38.0	438	4	US-09-252-991A-6582	Sequence 6582, Ap
C	38	22.8	38.0	1350	4	US-09-252-991A-6609	Sequence 6609, Ap
C	39	22.8	38.0	1482	4	US-09-252-991A-6692	Sequence 6692, Ap
C	40	22.8	38.0	1644	4	US-09-252-991A-6563	Sequence 6563, Ap
C	41	22.8	38.0	1791	4	US-09-252-991A-6662	Sequence 6662, Ap
C	42	22.6	37.7	1698	4	US-09-252-991A-887	Sequence 887, App
C	43	22.6	37.7	3889	4	US-09-484-970B-39	Sequence 39, Appl
C	44	22.6	37.7	33529	3	US-09-144-085-3	Sequence 3, Appl
C	45	22.6	37.7	4403765	3	US-09-103-840A-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-10-029-517-17
Sequence 17, Application US/10029517
Patent No. 6716627
GENERAL INFORMATION:
APPLICANT: Kenneth W. Dobie
TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
FILE REFERENCE: RTS-0352
CURRENT APPLICATION NUMBER: US/10/029,517
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 107
SEQ ID NO 17
LENGTH: 1804
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (73)...(1500)
US-10-029-517-17

Query Match 60.0%; Score 36; DB 4; Length 1804;
Best Local Similarity 75.0%; Pred. No. 0.0021;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTTCACGGCACCCTCCAGACAGGAGTACGCTGCACCCGACCCGTCAGTCCG 60
Db 457 GGCTCCACGGCCCCCGCCAGCCAGCGTGTACCTCGGCCCCCGACACAGCGGCCCG 516

RESULT 2
US-10-029-517-19
Sequence 19, Application US/10029517
Patent No. 6716627
GENERAL INFORMATION:
APPLICANT: Kenneth W. Dobie
TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
FILE REFERENCE: RTS-0352
CURRENT APPLICATION NUMBER: US/10/029,517
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 107
SEQ ID NO 19
LENGTH: 8186
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: unsure
LOCATION: 6899
OTHER INFORMATION: unknown

```

; NAME/KEY: unsure
; LOCATION: 7155
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7184
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7957
; OTHER INFORMATION: unknown
; NAME/KEY: intron
; LOCATION: (2997)...(3498)
; OTHER INFORMATION: intron 1
; NAME/KEY: intron:exon junction
; LOCATION: (3498)...(3499)
; OTHER INFORMATION: intron 1:exon 2
; NAME/KEY: exon
; LOCATION: (3508)...(3599)
; OTHER INFORMATION: exon 2d
; NAME/KEY: exon:intron junction
; LOCATION: (3982)...(3983)
; OTHER INFORMATION: exon 2a:intron 2a
; NAME/KEY: intron:exon junction
; LOCATION: (4205)...(4206)
; OTHER INFORMATION: intron 2c:exon 3c
; NAME/KEY: intron:exon junction
; LOCATION: (4259)...(4260)
; OTHER INFORMATION: intron 2d:exon 3d
; NAME/KEY: exon
; LOCATION: (4260)...(4328)
; OTHER INFORMATION: exon 3d
; NAME/KEY: intron:exon junction
; LOCATION: (4632)...(4633)
; OTHER INFORMATION: intron 3:exon 4
; NAME/KEY: exon
; LOCATION: (4914)...(5035)
; OTHER INFORMATION: exon 5
; NAME/KEY: intron
; LOCATION: (5266)...(6293)
; OTHER INFORMATION: intron 6
US-10-029-517-19
```

```
Query Match
Best Local Similarity 60.0%; Score 36; DB 4; Length 8186;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
```

```
QY 1 GGTTCACGGCAGCCTCCAGACGAGTCACTGTCACCCGACCCGCTCCAGCTCCG 60
Db 3825 GGCTCCACCGCCCCCCCCAGCCCGGTCCTCGGCCCCGAGACACGAGCCGCCCG 3884
```

```
RESULT 3
US-10-029-517-18
; Sequence 18, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 18
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (67)...(572)
US-10-029-517-18
```

```
Query Match
Best Local Similarity 57.3%; Score 34.4; DB 4; Length 572;
Matches 73.3%; Pred. No. 0.0059;
```

```
Matches 44; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 1 GGTTCACGGCAGCCTCCAGACGAGTCACTGTCACCCGACCCGCTCCAGCTCCG 60
Db 478 GGCTCCACCGCCCCCCCCAGCCCGGTCCTCGGCCCCCGAGACACGAGCCGCCCG 537
```

```
RESULT 4
US-10-029-517-3
; Sequence 3, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 3
; LENGTH: 1721
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (58)...(1605)
US-10-029-517-3
```

```
Query Match
Best Local Similarity 57.3%; Score 34.4; DB 4; Length 1721;
Matches 44; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
```

```
QY 1 GGTTCACGGCAGCCTCCAGACGAGTCACTGTCACCCGACCCGCTCCAGCTCCG 60
Db 442 GGCTCCACCGCCCCCCCCAGCCCGGTCCTCGGCCCCGAGACACGAGCCGCCCG 501
```

```
RESULT 5
US-10-029-517-101
; Sequence 101, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 101
; LENGTH: 518
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-10-029-517-101
```

```
Query Match
Best Local Similarity 56.0%; Score 33.6; DB 4; Length 518;
Matches 42; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
```

```
QY 1 GGTTCACGGCAGCCTCCAGACGAGTCACTGTCACCCGACCCGCTCCAGC 56
Db 462 GGCTCCACCGCCCCCCCCAGCCCGGTCCTCGGCCCCGAGACACGAGCCGCCG 517
```

```
RESULT 6
US-09-475-947A-246
; Sequence 246, Application US/09475947A
; Patent No. 6472154
; GENERAL INFORMATION:
; APPLICANT: Garner, Harold R.
; APPLICANT: Wren, Jonathan D.
; APPLICANT: Milna, John D.
```

```

; TITLE OF INVENTION: Polymorphic Repeats in Human Genes
; FILE REFERENCE: UTS0667
; CURRENT APPLICATION NUMBER: US/09/475,947A
; CURRENT FILING DATE: 1999-12-31
; NUMBER OF SEQ ID NOS: 346
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 246
; LENGTH: 60
; TYPE: DNA
; ORGANISM: human
US-09-475-947A-246

```

```

Query Match          53.7%; Score 32.2; DB 4; Length 60;
Best Local Similarity 75.5%; Pred. No. 0.022;
Matches 40; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```

```

QY      1 GGTTCACGGCAGCCTCCAGCAGCAGGAGTCAGCTGTCAGCCCGACCCGCTCC 53
Db      7 GGCTCCACCGCCCCCGCCAGCCAGCGGTGTCACTCGGCCCGGACACAGCGCC 59

```

RESULT 7

```

US-10-029-517-102
; Sequence 102, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myer
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 102
; LENGTH: 3343
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-10-029-517-102

```

```

Query Match          52.0%; Score 31.2; DB 4; Length 3343;
Best Local Similarity 70.0%; Pred. No. 0.1;
Matches 42; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```

```

QY      1 GGTTCACGGCAGCCTCCAGCAGCAGGAGTCAGCTGTCAGCCCGACCCGCTCCAGCTCCG 60
Db      1608 GGCTCCACCGCCCCCGCCAGCCAGATGTCACTCAGCCCGGAGACACAGCCAGCGCCG 1667

```

RESULT 8

```

US-08-479-537A-1
; Sequence 1, Application US/08479537A
; Patent No. 5861381
; GENERAL INFORMATION:
; APPLICANT: CHAMBON, Pierre
; APPLICANT: KIENY, Marie-Paule
; APPLICANT: LATHE, Richard
; APPLICANT: HAREUVENI, Mara
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
; TITLE OF INVENTION: TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

CURRENT APPLICATION DATA:

```

; APPLICATION NUMBER: US/08/479,537A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 90/13101
; FILING DATE: 23-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR91/00835
; FILING DATE: 23-OCT-1991
; APPLICATION NUMBER: US 08/039,320
; FILING DATE: 04-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,576
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6192 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 58..120
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 439..5239
; OTHER INFORMATION: /note= "The nucleotides spanning
; OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
; OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
; OTHER INFORMATION: The number of such repeats varies from 1 to 80."
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 121..6166
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 457
; OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 487
; OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
; OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
; OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
; NAME/KEY: repeat_region
; LOCATION: 496
; OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-08-479-537A-1

```

```

Query Match          52.0%; Score 31.2; DB 2; Length 6192;
Best Local Similarity 70.0%; Pred. No. 0.11;
Matches 42; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```

```

QY      1 GGTTCACGGCAGCCTCCAGCAGCAGGAGTCAGCTGTCAGCCCGACCCGCTCCAGCTCCG 60
Db      382 GGCTCCACCGCCCCCGCCAGCCAGATGTCACTCAGCCCGGAGCAACAAGCCAGCGCCG 441

```

RESULT 9

US-09-083-116-1

```
; Sequence 1, Application US/09083116
; Patent No. 6203795
; GENERAL INFORMATION:
; APPLICANT: CHAMBER, Pierre
; APPLICANT: KIENY, Marie-Paule
; APPLICANT: LATHE, Richard
; APPLICANT: HAREUVENI, Mara
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
; TITEL OF INVENTION: TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/083,116
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,537
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR91/00835
; FILING DATE: 23-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/039,320
; FILING DATE: 04-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,576
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 017753-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6192 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 58..120
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 439..5239
; OTHER INFORMATION: /note= "The nucleotides spanning
; OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
; OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
; FEATURE: The number of such repeats varies from 1 to 80."
; NAME/KEY: mat_peptide
; LOCATION: 121..6166
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 457
; OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
```

```
; NAME/KEY: repeat_region
; LOCATION: 487
; OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
; OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
; OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 496
; OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; US-09-083-116-1

Query Match 52.0%; Score 31.2; DB 3; Length 6192;
Best Local Similarity 70.0%; Pred. No. 0.11;
Matches 42; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCTCCAGCAGCAGGATGCTGTGACCCCGACACCCGTCAGTCGCG 60
DB 382 GGCTCCACCAACCCCGCCAGCCGACGATGTCACCTCAGCCCCGGACACACGACGCCCCG 441

RESULT 10
US-09-134-916A-1
; Sequence 1, Application US/09134916A
; Patent No. 6328956
; GENERAL INFORMATION:
; APPLICANT: CHAMBER, Pierre
; APPLICANT: KIENY, Marie-Paule
; APPLICANT: LATHE, Richard
; APPLICANT: HAREUVENI, Mara
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
; TITEL OF INVENTION: TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/134,916A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,537
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: FR 90/13101
; FILING DATE: 23-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR91/00835
; FILING DATE: 23-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/039,320
; FILING DATE: 04-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,576
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 017753-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
```

```

; LENGTH: 6192 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 58..120
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 439..5239
; OTHER INFORMATION: /note= "The nucleotides spanning
; OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
; OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
; OTHER INFORMATION: The number of such repeats varies from 1 to 80."
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 121..6166
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 457
; OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 487
; OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
; OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
; OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 496
; OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; US-09-134-916A-1

Query Match          52.0%; Score 31.2; DB 3; Length 6192;
Best Local Similarity 70.0%; Pred. No. 0.11;
Matches 42; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY      1 GGTCAACGGCACCCTCCAGACACGAGTCAGTGTGACCCGACACACCCGTCAGCTCCG 60
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      382 GGCTCCACCAACCCCGCCGACGATGTCACTCAGCCCGGACACAGCCAGCCCGG 441

RESULT 11
US-08-479-537A-4
; Sequence 4, Application US/08479537A
; Patent No. 5861381
; GENERAL INFORMATION:
; APPLICANT: CHAMON, Pierre
; APPLICANT: KIENY, Marie-Paule
; APPLICANT: LATHE, Richard
; APPLICANT: HAREUVENI, Mara
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
; TITLE OF INVENTION: TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,537A
```

```

; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 90/13101
; FILING DATE: 23-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR91/00835
; FILING DATE: 23-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/039,320
; FILING DATE: 04-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,576
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 017753-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6449 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 58..120
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 439..5239
; OTHER INFORMATION: /note= "The nucleotides spanning
; OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
; OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
; OTHER INFORMATION: The number of such repeats varies from 1 to 80."
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 121..5661
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 457
; OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCG
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 487
; OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
; OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACG
; OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 496
; OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCG
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; US-08-479-537A-4

Query Match          52.0%; Score 31.2; DB 2; Length 6449;
Best Local Similarity 70.0%; Pred. No. 0.11;
Matches 42; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY      1 GGTCAACGGCACCCTCCAGACACGAGTCAGTGTGACCCGACACACCCGTCAGCTCCG 60
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      382 GGCTCCACCAACCCCGCCGACGATGTCACTCAGCCCGGACACAGCCAGCCCGG 441

RESULT 12
US-09-083-116-4
; Sequence 4, Application US/09083116
```

Patent No. 6203795
GENERAL INFORMATION:
APPLICANT: CHAMBERON, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREUVENI, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/083,116
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,537
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6449 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note= "The nucleotides spanning
OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
FEATURE: The number of such repeats varies from 1 to 80."
NAME/KEY: mat_peptide
LOCATION: 121..5661
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
FEATURE: or CCG, and Ala = GCT, GCC, GCA, or GCG."
NAME/KEY: repeat_region
LOCATION: 487

OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG, and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG, and Ala = GCT, GCC, GCA, or GCG."
US-09-083-116-4
Query Match 52.0%; Score 31.2; DB 3; Length 6449;
Best Local Similarity 70.0%; Pred. No. 0.11;
Matches 42; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY 1 GGTCAACGGACCTCCAGCAGCAGGATGCTGCTGACCCGACACCCGTCAGCTCCG 60
Db 382 GGTCACACACCCCGCCAGCCGATGTCACCTCAGCCCGGACACAGCCAGCCCG 441
RESULT 13
US-09-134-916A-4
Sequence 4, Application US/09134916A
Patent No. 6328956
GENERAL INFORMATION:
APPLICANT: CHAMBERON, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREUVENI, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/134,916A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/479,537
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: FR 90/13101
FILING DATE: 23-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6449 base pairs
TYPE: nucleic acid


```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 58..120
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 439..5239
; OTHER INFORMATION: /note= "The nucleotides spanning
; OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
; OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
; OTHER INFORMATION: The number of such repeats varies from 1 to 80."
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 121..5661
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 457
; OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 487
; OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
; OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
; OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 496
; OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
;
US-09-134-916A-4

```

Query Match	52.0%;	Score 31.2;	DB 3;	Length 6449;
Best Local Similarity	70.0%;	Pred. No. 0.11;		
Matches 42;	Conservative 0;	Mismatches 18;	Indels 0;	Gaps 0;
Qy	1	GGTTCACGGGACCTCCAGCACA	CGAGTCACTGTGCACCCGACACCCGTCAGCTCCG	60
Db	382	GGCTCCACCACTCCCGCCAGCCGACGATGTCTACCTCAGCCCCCGACCAACAGCCAGCCCCG		441

```

RESULT 14
US-10-029-517-16
; Sequence 16, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 16
; LENGTH: 981
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon:exon junction
; LOCATION: (464)..(465)
; OTHER INFORMATION: exon 3b:exon 4
US-10-029-517-16

```

Query Match	50.7%;	Score 30.4;	DB 4;	Length 981;
Best Local Similarity	71.4%;	Pred. No. 0.15;		
Matches 40;	Conservative 0;	Mismatches 16;	Indels 0;	Gaps 0;

[illegible]

```

Db          21  GGCTCCACCGCCCCCGAGCCCATGTGTCTCACTCGGCCCGGACACAGGCGCCG 76

RESULT 15
US-09-646-028-42
; Sequence 42, Application US/09646028
; Patent No. 6562347
; GENERAL INFORMATION:
; APPLICANT: Kwak, Larry
; APPLICANT: Biragyn, Arya
; TITLE OF INVENTION: METHODS AND COMPOSITIONS OF
; TITLE OF INVENTION: CHEMOKINE-TUMOR ANTIGEN FUSION PROTEINS AS CANCER VACCINES
; FILE REFERENCE: 14014.0316/P
; CURRENT APPLICATION NUMBER: US/09/646,028
; CURRENT FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: 60/077,745
; PRIOR FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 42
; LENGTH: 519
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of artificial sequence:/note=synthetic construct
US-09-646-028-42

Query Match          47.7%; Score 28.6; DB 4; Length 519;
Best Local Similarity 67.8%; Pred. No. 0.56;
Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```

Oy 1 GGTTCACCGGCACTCCAGACACAGGAGTACGCTGTGCACCCGACACCCGTCACGCTCC 59
 Db 331 GGTTCACCTGCTCCGCGGACACAGGTGTAACTTCTGCGGCCCCGACGCGTAACTTC 389

Search completed: January 16, 2005, 03:01:51
Job time : 42.2 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 19:43:55 ; Search time 1550.6 seconds
(without alignments)
1410.024 Million cell updates/sec

Title: US-10-057-136-11

Perfect score: 60
Sequence: 1 ggtcaacgacacccaccacgc.....ccgacacccgctccagctccg 60

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37.6	62.7	604	4 BM791359	BM791359 K-EST0071
2	37.6	62.7	877	5 BUI542454	BUI542454 AGENCOURT
3	37.6	62.7	959	6 CA489836	CA489836 AGENCOURT
4	37.6	62.7	1113	5 BUI48487	BUI48487 AGENCOURT
5	37.6	62.7	1234	5 BQ936898	BQ936898 AGENCOURT
6	37.6	62.7	1262	5 BQ935496	BQ935496 AGENCOURT
7	37.6	62.7	1343	5 BQ920055	BQ920055 AGENCOURT
8	37.6	62.7	1531	5 BUI543309	BUI543309 AGENCOURT
9	36	60.0	330	1 AI925867	AI925867 w020d04.x
10	36	60.0	1130	5 BUI542996	BUI542996 AGENCOURT
11	36	60.0	1268	5 BQ943554	BQ943554 AGENCOURT
12	36	60.0	1349	5 BUI52566	BUI52566 AGENCOURT
13	36	60.0	1420	5 BUI542790	BUI542790 AGENCOURT
14	31.2	52.0	472	4 BM759495	BM759495 K-EST0039
15	30.4	50.7	981	4 BG774910	BG774910 602649832
16	30.4	50.7	1334	5 BQ943809	BQ943809 AGENCOURT
17	30.4	50.7	1536	5 BQ923149	BQ923149 AGENCOURT
18	29.6	49.3	1678	4 BG775565	BG775565 602650481
19	29.4	49.0	166	7 T27692	T27692 EST12384 Hu
20	29.4	49.0	475	6 CB120860	CB120860 K-EST0168
21	29.4	49.0	669	6 CB122585	CB122585 K-EST0170
22	29	48.3	619	4 B1260921	B1260921 602970962
23	28.8	48.0	932	4 BG444776	BG444776 GA_Ea002
24	28.6	47.7	1279	9 AG097186	AG097186 Pan trogl

C	25	27.6	46.0	356	8 AQ906176	AQ906176 GSSTC0667
C	26	27.6	46.0	395	4 BM098904	BM098904 EBP105_SQ
	27	27.4	45.7	540	1 AJ470277	AJ470277 AJ470277
	28	27.2	45.3	528	4 BM130288	BM130288 pb29b06.y
	29	27.2	45.3	538	4 B1744470	B1744470 pb09c12.y
	30	27.2	45.3	547	8 AQ841698	AQ841698 T137709b
	31	27.2	45.3	739	9 AG517386	AG517386 Mus muscu
	32	27.2	45.3	762	5 BX875037	BX875037 BX875037
C	33	27.2	45.3	768	9 CC901252	CC901252 ZMMBB023
C	34	27.2	45.3	787	9 CC877595	CC877595 ZMMBB021
	35	27.2	45.3	893	4 BM045068	BM045068 603622933
	36	27.2	45.3	1231	5 BQ890022	BQ890022 AGENCOURT
C	37	27	45.0	240	1 AJ460645	AJ460645 AJ460645
	38	27	45.0	262	1 AV919837	AV919837 AV919837
C	39	27	45.0	300	1 AJ460651	AJ460651 AJ460651
C	40	27	45.0	313	4 B1776854	B1776854 EBP103_SQ
C	41	27	45.0	314	8 AZ227363	AZ227363 RPT-23-7
C	42	27	45.0	315	4 BM097405	BM097405 EBem07_SQ
C	43	27	45.0	317	6 CB862996	CB862996 HH01P08u
C	44	27	45.0	323	6 CA028771	CA028771 H263D10x
	45	27	45.0	324	6 CB876471	CB876471 HX11G01w

ALIGNMENTS

RESULT 1
BM791359
LOCUS
DEFINITION K-EST0071342 S21SNUS20 Homo sapiens cDNA clone S21SNUS20-14-A06 5',
mRNA sequence.
ACCESSION BM791359 604 bp
VERSION BM791359.1 GI:19139591
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 604)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
TITLE Unpublished (2002)
JOURNAL Contact: Kim YS
COMMENT Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@email.kribb.re.kr
Plate: 14 row: A column: 06
High quality sequence stop: 604.
Location/Qualifiers
1. 604
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S21SNUS20-14-A06"
/sex="F"
/tissue_type="Stomach"
/cell_type="Floating aggregates"
/cell_line="SNU-520"
/lab_host="Top10F"
/clone_lib="S21SNUS20"
/note="Organ: Stomach; Vector: pTZ19RP1; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then deacapped
with tobacco acid pyrophosphatase (TAP). The deacapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dT-selected mRNA by
priming with dT-tailed vector. The dT-tailed vector was

adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for electroporation of competent cells E. coli Top10F⁺ by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

Query Match 62.7%; Score 37.6; DB 4; Length 604;
Best Local Similarity 76.7%; Pred. No. 0.12;
Matches 46; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCTCCAGACACGGAGTCAGCTTCGACCCGACACCCGTCAGCTCCG 60
DB 28 GGCTCCACCGCAGCCCGCCAGCCCGGTCACCTCGGCCCGGACACAGCGCCGCCCG 87

RESULT 2
BUS42454
LOCUS

DEFINITION BUS42454 877 bp mRNA linear EST 13-SEP-2002
AGENCOURT_10322173 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574322
5', mRNA sequence.

ACCESSION BUS42454
VERSION BUS42454.1 GI:22852937
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM2769 row: h column: 02
High quality sequence stop: 760.
Location/Qualifiers
1. .877

FEATURES
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6574322"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 62.7%; Score 37.6; DB 5; Length 877;
Best Local Similarity 76.7%; Pred. No. 0.12;
Matches 46; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCTCCAGCAGCAGGAGTCAGCTTCGACCCGACACCCGTCAGCTCCG 60
DB 215 GGCTCCACCGCAGCCCGCCAGCCCGGTCACCTCGGCCCGGACACAGCGCCGCCCG 274

RESULT 3
CA489836
LOCUS

DEFINITION CA489836 959 bp mRNA linear EST 14-NOV-2002
AGENCOURT_10810668 MAPCL Homo sapiens cDNA clone IMAGE:6722324 5',
mRNA sequence.

ACCESSION CA489836
VERSION CA489836
KEYWORDS CA489836.1 GI:24952627
SOURCE EST.
ORGANISM Homo sapiens (human)
Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Kristi A. Eglund, Ira Pastan
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM14284 row: n column: 20
High quality sequence start: 31
High quality sequence stop: 446.
Location/Qualifiers
1. .959

FEATURES
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6722324"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231,
hTERT-HME1, LNCap"
/lab_host="EMDH10B"
/clone_lib="MAPCL"
/note="Vector: pCMV-SPORT6; Site_1: EcoRV; Site_2: Not I;
Subtracted with brain, liver, lung, kidney and muscle.
directionally cloned. Priming method: oligo-dT. Average
insert size: 1800 bp. Library amplification: 26,000 fold.
Kristi A. Eglund, James J. Vincent, Robert Strausberg,
Bungkook Lee & Ira Pastan: Discovery of new breast
cancer genes encoding membrane and secreted proteins.
Manuscript submitted."

ORIGIN

Query Match 62.7%; Score 37.6; DB 6; Length 959;
Best Local Similarity 76.7%; Pred. No. 0.12;
Matches 46; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCTCCAGCAGCAGGAGTCAGCTTCGACCCGACACCCGTCAGCTCCG 60
DB 469 GGCTCCACCGCAGCCCGCCAGCCCGGTCACCTCGGCCCGGACACAGCGCCGCCCG 528

RESULT 4
BU148487
LOCUS

DEFINITION BU148487 1113 bp mRNA linear EST 03-SEP-2002
AGENCOURT_8670479 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6380642
5', mRNA sequence.

ACCESSION BU148487
VERSION BU148487.1 GI:22662019
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)

QY 1 GGTTCACGGCAGCTCCAGACACGAGTCTGACACCCGACCCGCTCCAGCTCCG 60
|||
Db 56 GGCTCCACCGCCCCCCCCAGCCGCGGTCTCAGCTTGCCCCCTGACACACGAGCGGCCCG 115
|||
RESULT 7
BQ920055 1343 bp mRNA linear EST 20-AUG-2002
LOCUS AGENCOURT_10031674 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6481609
5', mRNA sequence.
ACCESSION BQ920055
VERSION BQ920055.1 GI:22334753
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1343)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM2663 row: e column: 02
High quality sequence start: 56
High quality sequence stop: 237.
Location/Qualifiers
1. 1343
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6481609"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

Query Match 62.7%; Score 37.6; DB 5; Length 1343;
Best Local Similarity 76.7%; Pred. No. 0.12;
Matches 46; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 1 GGTTCACGGCAGCTCCAGACACGAGTCTGACACCCGACCCGCTCCAGCTCCG 60
|||
Db 72 GGCTCCACCGCCCCCCCCAGCCGCGGTCTCAGCTTGCCCCCTGACACACGAGCGGCCCG 131
|||
RESULT 8
BQ543309 1531 bp mRNA linear EST 13-SEP-2002
LOCUS AGENCOURT_10327072 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6575349
5', mRNA sequence.
ACCESSION BQ543309
VERSION BQ543309.1 GI:22853792
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 1531)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM2772 row: b column: 21
High quality sequence stop: 166.
Location/Qualifiers
1. 1531
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6575349"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

Query Match 62.7%; Score 37.6; DB 5; Length 1531;
Best Local Similarity 76.7%; Pred. No. 0.12;
Matches 46; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 1 GGTTCACGGCAGCTCCAGACACGAGTCTGACACCCGACCCGCTCCAGCTCCG 60
|||
Db 30 GGCTCCACCGCCCCCCCCAGCCGCGGTCTCAGCTTGCCCCCTGACACACGAGCGGCCCG 89
|||
RESULT 9
A1925867/c 330 bp mRNA linear EST 08-MAR-2000
LOCUS wo20d04.x1 NCI CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2455879 3'
DEFINITION similar to contains element MSRI MSRI repetitive element ;, mRNA
sequence.
ACCESSION A1925867
VERSION A1925867.1 GI:5661831
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 330)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert length: 578 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 305.
Location/Qualifiers
1. 330
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6575349"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 16:00:11 ; Search time 170.5 Seconds
(without alignments)
1847.304 Million cell updates/sec

Title: US-10-057-136-11

Perfect score: 60
Sequence: 1 ggttcacgacgcactccagc.....ccgacaccgcgtccagtcgcg 60

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapect 1.0

Searched: 4134886 segs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	2	AAV48324
C 2	40.8	68.0	156	10	ADK68635
3	39.2	65.3	1800	10	ADK68629
4	38.2	63.7	60	2	AAV48318
5	36.6	61.0	60	2	AAV48319
6	36	60.0	60	2	AAV48316
7	36	60.0	78	2	AAV48326
C 8	36	60.0	309	1	AAV48329
9	36	60.0	1194	12	AD157712
10	36	60.0	1378	12	AD157693
C 11	36	60.0	1424	12	ADO23180
12	36	60.0	1428	6	ABL60159
13	36	60.0	1428	12	ADO23125
14	36	60.0	1457	12	ADF32627
15	36	60.0	1527	2	AAV48329
16	36	60.0	1614	12	ADK70370
17	36	60.0	1630	12	AD157708
18	36	60.0	1634	12	AD157689
19	36	60.0	1712	12	AD157686
20	36	60.0	1738	12	AD157669
21	36	60.0	1755	12	AD157673

22	36	60.0	1774	12	ADK43991	Adk43991 Plasmid J
23	36	60.0	1774	12	ADF32625	Adf32625 Plasmid J
24	36	60.0	1799	12	ADO23124	Ado23124 Human MUC
25	36	60.0	1803	12	AD157699	Ad157699 Human bre
26	36	60.0	1804	6	ABL67539	Ab167539 Thyroid c
27	36	60.0	1804	9	AAD56950	Aad56950 Human muc
28	36	60.0	1804	10	ADD14719	Add14719 Human src
29	36	60.0	1804	12	ADP13294	Adp13294 Renal cel
30	36	60.0	1805	12	ADO28642	Ado28642 Human MUC
31	36	60.0	1808	12	AD157706	Ad157706 Human bre
32	36	60.0	1818	12	ADF32633	Adf32633 Plasmid J
33	36	60.0	1823	6	ABZ35228	Abz35228 Human gen
34	36	60.0	1823	12	AD157707	Ad157707 Human bre
35	36	60.0	1835	12	ADF32631	Adf32631 Plasmid J
36	36	60.0	1874	12	AD157688	Ad157688 Human bre
37	36	60.0	1882	12	AD157677	Ad157677 Human bre
38	36	60.0	1918	12	AD157672	Ad157672 Human bre
39	36	60.0	1930	12	AD157678	Ad157678 Human bre
40	36	60.0	1945	12	AD157676	Ad157676 Human bre
41	36	60.0	1949	12	AD157698	Ad157698 Human bre
42	36	60.0	1953	12	AD157668	Ad157668 Human bre
43	36	60.0	2045	12	AD157701	Ad157701 Human bre
44	36	60.0	2049	12	AD157682	Ad157682 Human bre
45	36	60.0	2090	12	AD157705	Ad157705 Human bre

ALIGNMENTS

RESULT 1
ID AAV48324
AAV48324 standard; DNA; 60 BP.

XX
AC AAV48324;
DT 20-NOV-1998 (first entry)
XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit R9.
XX
KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KM tumour; tumour-associated antigen.
XX
OS Homo sapiens.
XX
FN WO9837095-A2.
XX
PD 27-AUG-1998.
XX
PF 24-FEB-1998; 98WO-US003693.
XX
PR 24-FEB-1997; 97US-0038253P.
XX
PA (THER-) THERION BIOLOGICS CORP.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (DAND) DANA FARBER CANCER INST INC.
PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX
DR WPI; 1998-467492/40.
XX
PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS Disclosure; Page 11; 42pp; English.
XX
CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens

```
XX
SQ Sequence 60 BP; 12 A; 26 C; 14 G; 8 T; 0 U; 0 Other;

Query Match
Best Local Similarity 100.0%; Score 60; DB 2; Length 60;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCCTCCAGACACGAGTCAAGTCTGACCCGACACCCGCTCCAGCTCCG 60
    |||||
DB 1 GGTTCACGGCAGCCTCCAGACACGAGTCAAGTCTGACCCGACACCCGCTCCAGCTCCG 60

RESULT 2
ADK68635/C
ID ADK68635 standard; cDNA; 156 BP.
XX
AC ADK68635;
XX
DT 06-MAY-2004 (first entry)
XX
DE HSP65-MUC1 antigen CTL epitope related cDNA #3.
XX
KW Human breast cancer; Mycobacterium bovis; heat shock protein 65; BCG;
KW HSP65; cytotoxic T lymphocyte; CTL; MUC1; HSP65-MUC1 antigen CTL epitope;
KW ss.
XX
OS Unidentified.
XX
PN CN1368384-A.
XX
PD 11-SEP-2002.
XX
PF 08-FEB-2001; 2001CN-00102614.
XX
PR 08-FEB-2001; 2001CN-00102614.
XX
PA (DIWE-) DIWEIHUAYU BIO TECHNOLOGY CO LTD BEIJING.
XX
PI Yu Y, Li H;
XX
DR WPI; 2003-854662/80.
XX
PT Genetically engineered vaccine of MUC-1 antigen for human breast cancer.
XX
PS Example 2; Page 4 (Disclosure); 14pp; Chinese.
XX
CC The invention relates to a method of preparation of a genetically
CC engineered vaccine for preventing and treating human breast cancer. The
CC method comprises fusing the coding gene of the Mycobacterium bovis heat
CC shock protein 65 (BCG HSP65) with the cytotoxic T lymphocyte (CTL)
CC epitope gene of MUC1 (antigen cell expressed by human breast cancer
CC cells) and the HSP65-MUC1 antigen CTL epitope is expressed in Escherichia
CC coli cells. This sequence represents DNA used in the method of the
CC invention.
XX
SQ Sequence 156 BP; 37 A; 42 C; 58 G; 19 T; 0 U; 0 Other;

Query Match
Best Local Similarity 68.0%; Score 40.8; DB 10; Length 156;
Matches 48; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCCTCCAGACACGAGTCAAGTCTGACCCGACACCCGCTCCAGCTCCG 60
    |||||
DB 132 GGTTCACGGCAGCCTCCAGACACGAGTCAAGTCTGACCCGACACCCGCTCCAGCTCCG 73

RESULT 3
ADK68629
ID ADK68629 standard; cDNA; 1800 BP.
XX
AC ADK68629;
XX
DT 06-MAY-2004 (first entry)
XX
PA
```

```
XX
DE HSP65-MUC1 antigen CTL epitope related cDNA #1.
XX
KW Human breast cancer; Mycobacterium bovis; heat shock protein 65; BCG;
KW HSP65; cytotoxic T lymphocyte; CTL; MUC1; HSP65-MUC1 antigen CTL epitope;
KW gene; ss.
XX
OS Unidentified.
XX
PN CN1368384-A.
XX
PD 11-SEP-2002.
XX
PF 08-FEB-2001; 2001CN-00102614.
XX
PR 08-FEB-2001; 2001CN-00102614.
XX
PA (DIWE-) DIWEIHUAYU BIO TECHNOLOGY CO LTD BEIJING.
XX
PI Yu Y, Li H;
XX
DR WPI; 2003-854662/80.
XX
DR P-PSDB; ADK68630.
XX
PT Genetically engineered vaccine of MUC-1 antigen for human breast cancer.
XX
PS Claim 1; SEQ ID NO 1; 14pp; Chinese.
XX
CC The invention relates to a method of preparation of a genetically
CC engineered vaccine for preventing and treating human breast cancer. The
CC method comprises fusing the coding gene of the Mycobacterium bovis heat
CC shock protein 65 (BCG HSP65) with the cytotoxic T lymphocyte (CTL)
CC epitope gene of MUC1 (antigen cell expressed by human breast cancer
CC cells) and the HSP65-MUC1 antigen CTL epitope is expressed in Escherichia
CC coli cells. This sequence represents DNA used in the method of the
CC invention.
XX
SQ Sequence 1800 BP; 347 A; 561 C; 616 G; 276 T; 0 U; 0 Other;

Query Match
Best Local Similarity 65.3%; Score 39.2; DB 10; Length 1800;
Matches 47; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCCTCCAGACACGAGTCAAGTCTGACCCGACACCCGCTCCAGCTCCG 60
    |||||
DB 1639 GGTTCACGGCAGCCTCCAGACACGAGTCAAGTCTGACCCGACACCCGCTCCAGCTCCG 1698

RESULT 4
AAV48318
ID AAV48318 standard; DNA; 60 BP.
XX
AC AAV48318;
XX
DT 20-NOV-1998 (first entry)
XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit R3.
XX
KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW tumour; tumour-associated antigen.
XX
OS Homo sapiens.
XX
PN WO9837095-A2.
XX
PD 27-AUG-1998.
XX
PF 24-FEB-1998; 98WO-US003693.
XX
PR 24-FEB-1997; 97US-0038253P.
XX
PA (THER-) THERION BIOLOGICS CORP.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
```



```
Db      1 GGCTCCACCGCCCCCAGCCACGGTGTCACTCGCCCCCGAGACCAAGCGGCCCCG 60
RESULT 7
AAV48326
ID      AAV48326 standard; cDNA; 78 BP.
XX
AC      AAV48326;
XX
DT      20-NOV-1998 (first entry)
XX
DE      Nucleotide sequence encoding MUC1 tandem repeat unit a.
XX
KM      ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KM      tumour; tumour-associated antigen.
XX
OS      Homo sapiens.
XX
FH      Key
FT      CDS
FT      1..78
FT      /tag= a
FT      /product= "MUC1 tandem repeat unit"
XX
PN      WO9837095-A2.
XX
PD      27-AUG-1998.
XX
PF      24-FEB-1998; 98WO-US003693.
XX
PR      24-FEB-1997; 97US-0038253P.
XX
PA      (THER-) THERION BIOLOGICS CORP.
PA      (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PA      (DAND) DANA FARBER CANCER INST INC.
XX
PI      Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX
DR      WPI; 1998-467492/40.
DR      P-PSDB; AAW77230.
XX
PT      New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT      immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS      Example 1; Page 20; 42pp; English.
XX
CC      The MUC1 tandem repeat units AAV48326-V48328 were used to create an
CC      immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC      (RPV). The RPV was used in a pharmaceutical composition also containing
CC      an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC      recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC      does not undergo significant genetic deletion, thereby providing an
CC      unexpectedly stable and immunogenic pox virus. They can be used to
CC      prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ      Sequence 78 BP; 9 A; 42 C; 22 G; 5 T; 0 U; 0 Other;
XX
Query Match
Best Local Similarity 60.0%; Score 36; DB 2; Length 78;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
OY      1 GGTTCACGGCAGCTCCAGACACGAGTCAAGTGTGACACCCGACACCCGTCAGCTCCG 60
Db      1 GGCTCCACCGCAGCCCGACCGCCACGGTGTCACTCGCCCCCGAGACCAAGCGGCCCCG 60
RESULT 8
AAAN90579/c
ID      AAAN90579 standard; cDNA; 309 BP.
XX
AC      AAAN90579;
XX
DT      27-AUG-2003 (revised)
DT      25-MAR-2003 (revised)
XX
```

```
DT      04-DEC-1989 (first entry)
XX
DE      PDF9.3 cDNA insert.
XX
KM      PDF9.3; human DF3 breast carcinoma-associated antigen epitope.
XX
OS      Human MCF-7 breast carcinoma cells; 'ATCC HTB22'.
XX
PN      WO8907107-A.
XX
PD      10-AUG-1989.
XX
PF      29-JAN-1988; 88US-00149831.
XX
PR      29-JAN-1988; 88US-00149831.
XX
PA      (DANA-) DANA-FARBER CANCER.
XX
PI      Kufe DW;
XX
DR      WPI; 1989-248989/34.
DR      P-PSDB; AAP91045, AAP91053, AAP91054, AAP90146.
XX
PT      Recombinant polypeptide(s) - contains DF 3 breast carcinoma antigen
PT      epitope and useful as assay reagents, and encoding DNA sequences.
XX
PS      Claim 1; Fig 4; 31pp; English.
XX
CC      The sequence encodes a carbohydrate-free polypeptide contg. DF3 breast
CC      carcinoma antigen epitope. Useful as a competitive binding assay reagent
CC      and improves diagnosis. The cDNA consists of nearly identical 60 BP
CC      tandem repeats which are 85% GC-rich. See also AAP91045, AAP91046,
CC      AAP91053, AAP91054. (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ      Sequence 309 BP; 21 A; 92 C; 160 G; 36 T; 0 U; 0 Other;
XX
Query Match
Best Local Similarity 60.0%; Score 36; DB 1; Length 309;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
OY      1 GGTTCACGGCAGCTCCAGACACGAGTCAAGTGTGACACCCGACACCCGTCAGCTCCG 60
Db      266 GGCTCCACCGCCCCCAGCCACGGTGTCACTCGCCCCCGAGACCAAGCGGCCCCG 207
RESULT 9
AD157712
ID      AD157712 standard; cDNA; 1194 BP.
XX
AC      AD157712;
XX
DT      22-APR-2004 (first entry)
XX
DE      Human breast specific nucleic acid (BSNA) #83.
XX
KW      Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
KW      breast cancer; cytostatic.
XX
OS      Homo sapiens.
XX
PN      WO2003106648-A2.
XX
PD      24-DEC-2003.
XX
PF      16-JUN-2003; 2003WO-US018934.
XX
PR      14-JUN-2002; 2002US-0389327P.
XX
PA      (DIAD-) DIADEXUS INC.
XX
PI      Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
XX
```


DR WPI; 2004-082185/08.
XX P-PSDB; ADI57782.
PT Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.
XX
XX Claim 1; SEQ ID NO 83; 370pp; English.
XX
CC The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the
CC hybridisation. The nucleic acids are useful for determining the presence
CC of a BSP in a sample which involves contacting the sample with a BSP in
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein and
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.
XX
SQ Sequence 1194 BP; 243 A; 438 C; 275 G; 238 T; 0 U; 0 Other;
Query Match 60.0%; Score 36; DB 12; Length 1194;
Best Local Similarity 75.0%; Pred. No. 0.022;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 GGTTCACGGCACCCTCCAGCACACGAGTCACGTCTGCACCCGACACCCCGTCCAGCTCCG 60
Db 596 GGCTCCACCGCCCCCCCCAGCCACGGGTGTCACTCGGCCCCCGACACACGCGGCCCG 655
RESULT 10
AD157693
ID AD157693 standard; cDNA; 1378 BP.
XX
AC AD157693;
XX
DT 22-APR-2004 (first entry)
XX
DE Human breast specific nucleic acid (BSNA) #64.
XX
DE Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
KW breast cancer; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2003106648-A2.
XX
PD 24-DEC-2003.
XX
PF 16-JUN-2003; 2003WO-US018934.
XX
PR 14-JUN-2002; 2002US-0389327P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
XX
DR WPI; 2004-082185/08.
DR P-PSDB; ADI57765.
XX
PT Novel isolated polypeptide comprising breast specific protein sequences,

PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.
XX
XX Claim 1; SEQ ID NO 64; 370pp; English.
XX
CC The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the
CC hybridisation. The nucleic acids are useful for determining the presence
CC of a BSP in a sample which involves contacting the sample with a suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein and
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.
XX
SQ Sequence 1378 BP; 295 A; 489 C; 316 G; 278 T; 0 U; 0 Other;
Query Match 60.0%; Score 36; DB 12; Length 1378;
Best Local Similarity 75.0%; Pred. No. 0.022;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 GGTTCACGGCACCCTCCAGCACACGAGTCACGTCTGCACCCGACACCCCGTCCAGCTCCG 60
Db 596 GGCTCCACCGCCCCCCCCAGCCACGGGTGTCACTCGGCCCCCGACACACGCGGCCCG 655
RESULT 11
ADO23180/c.
ID ADO23180 standard; RNA; 1424 BP.
XX
AC ADO23180;
XX
DT 12-AUG-2004 (first entry)
XX
DE Antisense human MUC1 mucin glycoprotein RNA (coding sequence) SeqID 75.
XX
DE human; MUC1; mucin glycoprotein; cancer; chemotherapeutic; MUC1/ECG; ss;
KW antisense.
XX
OS Homo sapiens.
XX
PN WO2004044160-A2.
XX
PD 27-MAY-2004.
XX
PF 12-NOV-2003; 2003WO-US035848.
XX
PR 13-NOV-2002; 2002US-00293391.
XX
PR 29-MAY-2003; 2003US-00447839.
XX
PA (DAND) DANA FARBER CANCER INST INC.
XX
PA (ILEX-) ILEX PROD INC.
XX
PI Kufe DW, Kharbanda S, Weltman SD;
XX
DR WPI; 2004-420304/39.
XX
XX Double-stranded RNA complex useful for inhibiting proliferation of cancer
PT cell expressing MUC1 mucin glycoprotein, comprises first and second RNA
PT sequences.

PS Disclosure; SEQ ID NO 75; 112pp; English.

XX
CC This invention relates to novel modulators of the human MUC1 mucin
CC glycoprotein for use in cancer therapeutics, where MUC1 is a protein that
CC acts to inhibit the apoptotic response to genotoxic stress caused by
CC chemotherapeutic agents. In particular, it refers to modulators of the
CC MUC1 extracellular domain (MUC1/ECD). The method refers to using double-
CC stranded RNA complexes as MUC1 interference RNA compositions such that
CC MUC1 expression is inhibited, which in turn inhibits cancer cell
CC proliferation. The present invention describes screening assays to
CC identify compounds that inhibit the binding of various MUC1 ligands such
CC as neutregulin 2, as well as agonists, antagonists and antibodies thereof.
CC Furthermore, it provides MUC1 antisense, small interfering RNA (siRNA)
CC and small molecules in combination with chemotherapeutic agents that are
CC useful in the field of cancer therapy. This polynucleotide sequence is
CC the antisense human MUC1 RNA of the invention.

XX
SQ Sequence 1424 BP; 309 A; 325 C; 495 G; 0 T; 295 U; 0 Other;

Query Match 60.0%; Score 36; DB 12; Length 1424;
Best Local Similarity 75.0%; Pred. No. 0.022;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY 1 GGTCAACGGCACCCTCCAGACACGAGTCACGTCGACCCGACACCCGTCAGCTCCG 60
Db 1041 GGCTCCACCGCCCCCGCCAGCCACGCGTGTCACTCGGCCCCGACACAGCGCCGCCCG 982

RESULT 12
ABL60159

ID ABL60159 standard; cDNA; 1428 BP.

AC ABL60159;

DT 22-JUL-2002 (first entry)

DE Human MUC1 encoding cDNA SEQ ID NO 2.

XX
KW Human; mucin 1; MUC1; transmembrane protein; SNP; cancer; cytostatic;
KW single nucleotide polymorphism; haplotyping; genotyping; drug;
XX antinflammatory; gene; ss.
XX
OS Homo sapiens.

XX
FH Key Location/Qualifiers
FT CDS 1..1428

FT /*tag= a
FT /product= "MUC1"
FT replace(1009,A)
FT /*tag= b

FT /strand name= "Single nucleotide polymorphism"
FT /note= "SNP allelic variation results in Val substituted
FT by Met at position 337 of the MUC1 protein (ABB77476)"
XX

PN WO200226765-A2.

PD 04-APR-2002.

PF 25-SEP-2001; 2001WO-US030151.

PR 28-SEP-2000; 2000US-0236113P.

XX
PA (GENA-) GENAISANCE PHARM INC.

PI Chew A, Koshy B;

XX
DR WPI; 2002-405042/43.

DR P-PSDB; ABB77476.

XX
PT New genetic variants of mucin 1, Transmembrane gene, useful in studying
PT expression and function of protein encoded by the gene and for screening
PT drugs to treat diseases e.g. cancer.
XX

PS Claim 23; Fig 2; 75pp; English.

XX
CC The invention relates to a polynucleotide (ABL60158, ABL60159) encoding
CC mucin 1/MUC1 (ABB77476), Transmembrane isogene. The invention describes
CC novel genetic variants of the MUC1 gene. The invention is useful for
CC haplotyping/genotyping the MUC1 gene in an individual and identifying an
CC association between a trait and at least one of the haplotypes or
CC haplotype pairs of MUC1 gene. MUC1 is useful for studying the expression
CC and function of MUC1 and expressing MUC1 protein for use in screening for
CC candidate drugs to treat diseases related to MUC1 activity and in
CC studying the effect of the variation on the biological activity of MUC1
CC as well as on the binding affinity of candidate drugs targeting MUC1 for
CC the treatment of e.g. cancer. MUC1 is further used by the pharmaceutical
CC research scientist to validate MUC1 as a candidate target for and in
CC design of clinical trials of candidate drugs for, treating a specific
CC condition drugs or disease predicted to be associated with MUC1 activity.
CC MUC1 antibodies are useful in a variety of diagnostic and prognostic
CC formats and therapeutic methods

XX
SQ Sequence 1428 BP; 296 A; 496 C; 327 G; 309 T; 0 U; 0 Other;

Query Match 60.0%; Score 36; DB 6; Length 1428;
Best Local Similarity 75.0%; Pred. No. 0.022;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY 1 GGTCAACGGCACCCTCCAGACACGAGTCACGTCGACCCGACACCCGTCAGCTCCG 60
Db 385 GGCTCCACCGCCCCCGCCAGCCACGCGTGTCACTCGGCCCCGACACAGCGCCGCCCG 444

RESULT 13

ADO23125

ID ADO23125 standard; RNA; 1428 BP.

XX
AC ADO23125;

DT 12-AUG-2004 (first entry)

DE Human MUC1 mucin glycoprotein RNA (coding sequence) SeqID 20.

XX
KW human; MUC1; mucin glycoprotein; cancer; chemotherapeutic; MUC1/ECD; ss.

XX
OS Homo sapiens.

PN WO2004044160-A2.

PD 27-MAY-2004.

PF 12-NOV-2003; 2003WO-US035848.

PR 13-NOV-2002; 2002US-00293391.

PR 29-MAY-2003; 2003US-00447839.

XX
PA (DAND) DANA FARBER CANCER INST INC.

XX
PA (ILEX-) ILEX PROD INC.

PI Kufe DW, Kharbada S, Weltman SD;

XX
DR WPI; 2004-420304/39.

XX
PT Double-stranded RNA complex useful for inhibiting proliferation of cancer
PT cell expressing MUC1 mucin glycoprotein, comprises first and second RNA
PT sequences.
XX

PS Claim 2; SEQ ID NO 20; 112pp; English.

XX
CC This invention relates to novel modulators of the human MUC1 mucin
CC glycoprotein for use in cancer therapeutics, where MUC1 is a protein that
CC acts to inhibit the apoptotic response to genotoxic stress caused by
CC chemotherapeutic agents. In particular, it refers to modulators of the
CC MUC1 extracellular domain (MUC1/ECD). The method refers to using double-
CC stranded RNA complexes as MUC1 interference RNA compositions such that
CC MUC1 expression is inhibited, which in turn inhibits cancer cell

CC proliferation. The present invention describes screening assays to
CC identify compounds that inhibit the binding of various MUC1 ligands such
CC as neutregulin 2, as well as agonists, antagonists and antibodies thereof.
CC Furthermore, it provides MUC1 antisense, small interfering RNA (siRNA)
CC and small molecules in combination with chemotherapeutic agents that are
CC useful in the field of cancer therapy. This polynucleotide sequence is
CC the human MUC1 RNA of the invention.

XX
SQ Sequence 1428 BP; 296 A; 496 C; 327 G; 0 T; 309 U; 0 Other;

Query Match 60.0%; Score 36; DB 12; Length 1428;

Best Local Similarity 70.0%; Pred. No. 0.022;
Matches 42; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

OY 1 GGTCAACGGCACCTCCAGACACGAGTCAAGTGTGACACCCGTCAGTCCG 60
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 385 GGCTCCACCGCCCCCGCCAGCGGUGCACCGGCCCGGACACGAGCGGCCCG 444

RESULT 14

ADP32627
ID ADF32627 standard; DNA; 1457 BP.

XX
AC ADF32627;

XX
DT 26-FEB-2004 (first entry)

XX
DE Plasmid JNM358 MUC-1 nucleotide sequence.

XX
KM MUC-1 antigen; immune response; MUC-1; variable number of tandem repeat;
KW VNTR; repeat unit; tumour; metastasis; cytostatic; vaccine; gene therapy;
XX gene; ds.

XX
OS Synthetic.

XX
PN WO2003100060-A2.

XX
PD 04-DEC-2003.

XX
PF 23-MAY-2003; 2003WO-EP005594.

XX
PR 24-MAY-2002; 2002GB-00012046.

XX
PA (GLAXO) GLAXO GROUP LTD.

XX
PI Burden N, Ellis JH, Hamblin PA;

XX
DR WPI; 2004-042811/04.

PT New nucleic acid molecule encoding a MUC-1 antigen, useful for preparing
PT a composition for treating or preventing tumors or metastases.

XX
PS Example; Fig 3; 66pp; English.

XX
CC The present invention describes a nucleic acid molecule which encodes a
CC MUC-1 antigen. The nucleic acid is capable of raising an immune response
CC in vivo, has reduced susceptibility to recombination than full-length MUC
CC -1 and comprises between 1 and 15 variable number of tandem repeats
CC (VNTR) perfect repeat units. Also described: (1) a plasmid comprising the
CC DNA molecule; (2) a protein encoded by the nucleic acid; (3) a
CC pharmaceutical composition comprising the nucleic acid, plasmid or
CC protein and an excipient, diluent or carrier; and (4) a method of
CC treating or preventing tumors or metastases. A MUC1 antigen has
CC cytostatic activity, and can be used in vaccines, and in gene therapy.
CC The nucleic acid is useful for preparing a composition for treating or
CC preventing tumors or metastases. The present sequence is used in the
CC exemplification of the present invention.

XX
SQ Sequence 1457 BP; 302 A; 506 C; 334 G; 315 T; 0 U; 0 Other;

Query Match 60.0%; Score 36; DB 12; Length 1457;

Best Local Similarity 75.0%; Pred. No. 0.022;

Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY 1 GGTCAACGGCACCTCCAGACACGAGTCAAGTGTGACACCCGTCAGTCCG 60
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 403 GGCTCCACCGCCCCCGCCAGCGGUGTCACTCGGCCCGGACACGAGCGGCCCG 462

RESULT 15

AAV48329
ID AAV48329 standard; cDNA; 1527 BP.

XX
AC AAV48329;

XX
DT 20-NOV-1998 (first entry)

XX
DE MiniMUC1 gene.

XX
KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW tumour; tumour-associated antigen.

XX
OS Homo sapiens.

XX
FH Key Location/Qualifiers

FT CDS 1..1527

FT /tag= a
FT /product= "MiniMUC1 protein"

XX
PN WO9837095-A2.

XX
PD 27-AUG-1998.

XX
PF 24-FEB-1998; 98WO-US003693.

XX
PR 24-FEB-1997; 97US-0038253P.

XX
PA (THER-) THERION BIOLOGICS CORP.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (DAND) DANA FARBER CANCER INST INC.

XX
PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;

XX
DR WPI; 1998-467492/40.

XX
DR P-PSDB; AAW77233.

PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.

XX
PS Example 1; Page 21-22; 42pp; English.

XX
CC The immunogenic mini-MUC1 fragment was created from MUC1 tandem repeat
CC units for inclusion in a recombinant pox virus (RPV). The RPV was used in
CC a pharmaceutical composition also containing an immunomodulator to
CC generate MUC1 specific cytotoxic T-lymphocytes. The recombinant pox virus
CC therefore encodes an immunogenic MUC1 fragment that does not undergo
CC significant genetic deletion, thereby providing an unexpectedly stable
CC and immunogenic pox virus. They can be used to prevent or treat tumours
CC expressing MUC1 tumour-associated antigens

XX
SQ Sequence 1527 BP; 296 A; 573 C; 351 G; 307 T; 0 U; 0 Other;

Query Match 60.0%; Score 36; DB 2; Length 1527;
Best Local Similarity 75.0%; Pred. No. 0.023;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY 1 GGTCAACGGCACCTCCAGACACGAGTCAAGTGTGACACCCGTCAGTCCG 60
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 226 GGCTCCACCGCCCCCGCCAGCGGUGTCACTCGGCCCGGACACGAGCGGCCCG 285

Search completed: January 15, 2005, 20:36:11
Job time : 170.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 17:40:52 : Search time 722.8 Seconds
(without alignments)
3925.542 Million cell updates/sec

Title: US-10-057-136-11

Perfect score: 60
Sequence: 1 ggtcaacgacgcaccaccagc.....ccgacacccgctccagctccg 60

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	2297	6	BD272907 A recombi
2	36	60.0	120	6	AX192396 Sequence
3	36	60.0	120	6	BD000571 Human pol
4	36	60.0	1414	12	AF423031 Synthetic
5	36	60.0	1455	6	CQ715242 Sequence
6	36	60.0	1457	6	AX959914 Sequence
7	36	60.0	1774	6	AX959684 Sequence
8	36	60.0	1774	6	AX959912 Sequence
9	36	60.0	1804	6	AR492318 Sequence
10	36	60.0	1804	6	AX335367 Sequence
11	36	60.0	1804	9	HUMMUCAB
12	36	60.0	1834	12	AF423030 Synthetic
13	36	60.0	1835	6	AX959918 Sequence
14	36	60.0	2135	6	AX959916 Sequence
15	36	60.0	4139	6	CQ834017 Sequence
16	36	60.0	4139	6	AX334899 Sequence
17	36	60.0	4139	6	AX335372 Sequence
18	36	60.0	4139	6	AX336712 Sequence
19	36	60.0	4139	6	AX409474 Sequence

20	36	60.0	4139	6	AX440481 Sequence
21	36	60.0	4139	9	J05582 Human pancr
22	36	60.0	7188	9	AY463543 Homo sapi
23	36	60.0	8181	6	AX406624 Sequence
24	36	60.0	8181	9	M61170 Human polym
25	36	60.0	8186	6	AR492320 Sequence
26	36	60.0	133525	9	AL713399 Human DNA
27	36	57.3	120	6	AX192397 Sequence
28	34.4	57.3	120	6	BD000572 Human pol
29	34.4	57.3	180	6	E08763 CDNA encodi
30	34.4	57.3	572	6	AR492319 Sequence
31	34.4	57.3	572	9	HUMDF3AA
32	34.4	57.3	1721	6	CQ771290 Sequence
33	34.4	57.3	1721	6	AR492306 Sequence
34	34.4	57.3	1721	6	AX335860 Sequence
35	34.4	57.3	1721	6	AX440427 Sequence
36	34.4	57.3	1721	6	AX587588 Sequence
37	34.4	57.3	1721	9	HSTBYMA
38	34.4	57.3	1971	6	AX963157 Sequence
39	34.4	57.3	2037	6	AX963159 Sequence
40	34.4	57.3	2238	9	HSSETA
41	34.4	57.3	2751	9	AF176947 Macaca mu
42	33.6	56.0	491	9	HUMEPISB1
43	33.6	56.0	518	6	AR492402 Sequence
44	33.6	56.0	518	9	HUMEPIS1A1
45	33.6	56.0	1320	6	E08764 CDNA encodi

ALIGNMENTS

RESULT 1	BD272907	2297 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD272907	A recombinant vector expressing multiple constitutimulatory molecules and uses thereof.			
DEFINITION	BD272907	A recombinant vector expressing multiple constitutimulatory molecules and uses thereof.			
ACCESSION	BD272907				
VERSION	BD272907.1	GI:33082675			
KEYWORDS	JP 2002531133-A/1.				
SOURCE	synthetic construct				
ORGANISM	synthetic construct				
REFERENCE	artificial sequences.				
AUTHORS	1 (bases 1 to 2297)				
TITLE	Schlom,J., Hodge,J. and Panicali,D.				
JOURNAL	A recombinant vector expressing multiple constitutimulatory molecules and uses thereof				
COMMENT	Patent: JP 2002531133-A 1 24-SEP-2002;				
	THE UNITED STATES OF AMERICA, THERION BIOLOGICS CORP				
	OS Artificial Sequence				
	PN JP 2002531133-A/1				
	PD 24-SEP-2002				
	PF 12-NOV-1999 JP 2000586927				
	PR 09-DEC-1998 US 60/111582				
	PI JEFFREY SCHLOM, JAMES HODGE, DENNIS PANICALI				
	PC C12N15/02, A61K35/12, A61K35/74, A61K35/76, A61K38/00, A61K39/00,				
	PC A61K39/12,				
	PC A61K39/125, A61K39/21, A61K39/235, A61K39/245, A61K39/275, A61K39/				
	PC 29, A61K48/00,				
	PC A61P1/04, A61P31/04, A61P31/10, A61P31/12, A61P35/00, A61P37/02, PC				
	A61P37/06,				
	PC C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12N7/00, C12Q1/02, G01N33/				
	PC 53, C12N15/00,				
	PC C12N5/00, A61K37/02				
	CC Description of Artificial Sequence: VECTOR SEQUENCE FH	Key			
	FT Location/Qualifiers				
	FT source 1. .2297				
FEATURES	Location/Qualifiers	/organism='Artificial Sequence'.			
source	1. .2297				
	/organism="synthetic construct"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:32630"				
ORIGIN					

Query Match
Best Local Similarity 100.0%; Score 60; DB 6; Length 2297;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCACGGCACCCTCCAGCACACGAGTCAAGTCTTGACCCGACACCCGTCAGCTCCG 60
DB 826 GGTTCACGGCACCCTCCAGCACACGAGTCAAGTCTTGACCCGACACCCGTCAGCTCCG 885

RESULT 2
AX192396/c
LOCUS AX192396 120 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 2 from Patent EPI103623.
ACCESSION AX192396
VERSION AX192396.1 GI:15210363
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1
AUTHORS Taylor-Papadimitriou, J., Burchell, J. and Gendler, S.
TITLE Human mucin core protein: nucleic acid probes, peptide fragments and antibodies thereto, and uses thereof in diagnostic and therapeutic methods
JOURNAL Patent: EP 1103623-A 2 30-MAY-2001;
FEATURES IMPERIAL CANCER RESEARCH TECHNOLOGY LIMITED (GB)
source 1..120
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match
Best Local Similarity 75.0%; Score 36; DB 6; Length 120;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTTCACGGCACCCTCCAGCACACGAGTCAAGTCTTGACCCGACACCCGTCAGCTCCG 60
DB 87 GGTTCACGGCACCCTCCAGCACACGAGTCAAGTCTTGACCCGACACCCGTCAGCTCCG 28

RESULT 3
BD000571/c
LOCUS BD000571 120 bp DNA linear PAT 31-JAN-2002
DEFINITION Human polymorphic epithelial mucin core protein and nucleic acid encoding the protein.
ACCESSION BD000571
VERSION BD000571.1 GI:18623684
KEYWORDS JP 200033675-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1
AUTHORS Papadimitriou, J.T., Jendora, S. and Bachieru, J.
TITLE Human polymorphic epithelial mucin core protein and nucleic acid encoding the protein
JOURNAL Patent: JP 200033675-A 2 05-DEC-2000;
COMMENT IMPERIAL CANCER RESEARCH TECHNOLOGY LTD
OS Homo sapiens (human)
PN JP 200033675-A/2
PD 05-DEC-2000
PF 26-APR-2000 JP 2000125724
PR 07-JAN-1987 GB 8700269, 07-JAN-1987 GB 8700279 PR
22-APR-1987 US 041306, 09-NOV-1987 GB 8726172 FI JOYCE
TAYLOR PAPADIMITRIOU, SANDRA JENDORA, JOY BACHIERU PC
C12N15/02, A61K38/00, A61K39/395, A61K39/395, A61K49/00, A61P35/00, PC
C07K14/47,
PC C07K16/44, C12N5/10, C12P21/08// (C12N15/02, C12R1:91), (C12N5/10, C12R1:91),

PC C12N15/00, A61K37/02, C12N5/00, (C12N15/00, C12R1:91), (C12N5/00, C12R1:91)
PC C12R1:91)
CC
FH Key Location/Qualifiers
FT source 1..120
FT /organism="Homo sapiens (human)"
FEATURES
source 1..120
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match
Best Local Similarity 60.0%; Score 36; DB 6; Length 120;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTTCACGGCACCCTCCAGCACACGAGTCAAGTCTTGACCCGACACCCGTCAGCTCCG 60
DB 87 GGTTCACGGCACCCTCCAGCACACGAGTCAAGTCTTGACCCGACACCCGTCAGCTCCG 28

RESULT 4
AF423031
LOCUS AF423031 1414 bp mRNA linear SYN 10-JUL-2003
DEFINITION Synthetic construct Homo sapiens mucin variant MUC1-CT58 (MUC1) mRNA, complete cds; alternatively spliced.
ACCESSION AF423031
VERSION AF423031.1 GI:19338621
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1
AUTHORS Hinojosa-Kurtzberg, A.M., Johansson, M.E., Madsen, C.S., Hansson, G.C. and Gendler, S.J.
TITLE Novel MUC1 splice variants contribute to mucin overexpression in CTR-deficient mice
JOURNAL Am. J. Physiol. Gastrointest. Liver Physiol. 284 (5), G853-G862 (2003)
MEDLINE 22570517
PUBMED 12529261
REFERENCE 2 (bases 1 to 1414)
Hinojosa-Kurtzberg, A.M. and Gendler, S.J.
AUTHORS Direct Submission
TITLE Submitted (21-SEP-2001) Biochemistry, Mayo Clinic, 13400 E. Shea Boulevard, Scottsdale, AZ 85259, USA
JOURNAL
FEATURES
source
location/Qualifiers
1..1414
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/note="Homo sapiens gene in transgenic Mus musculus C57BL/6; isolated from intestinal mucosa"
1..1414
/gene="MUC1"
1..1386
/gene="MUC1"
/note="alternatively spliced; contains exon 6b resulting in variant carboxy-terminal domain; lacks sites for beta-catenin and Grb2 interactions; derived from Homo sapiens"
/codon_start=1
/transl_table=11
/product="mucin variant MUC1-CT58"
/protein_id="AA186735.1"
/db_xref="GI:19338622"
/translation="MTPTQSPFLLLLTLVTVTSGHASTPGGEKETSATORSS VPSSTKNAVSMTSSVLSHSPGSGSSTQGDVTLAPATEPASGSAATWGDVTSVP VTRPALGSTTPPAHDVTSAPDNKPAAGSTAPPAHGVTSPDRPAPGSTAPPAHGVTS APDNRPALGSTAPPVHNVTSAGSASGASASTLVHNGTSARATTPASSTPFSIPSHH SDPTTIAHSTKTVDASSTHSTVPLTSSNHSSTSPQLSTGVSFFLSFHSNLOFNS SLDEPSTDYQELQRDISEMFLQIYKGGFLGLSNIKFRPGSVVVLTLAIREGTINV

TITLE	Novel MUC1 splice variants contribute to mucin overexpression in CFTR-deficient mice
JOURNAL	Am. J. Physiol. Gastrointest. Liver Physiol. 284 (5), G853-G862 (2003)
MEDLINE	22570517
PUBMED	12529261
REFERENCE	2 (bases 1 to 1834)
AUTHORS	Hinojosa-Kurtzberg, A.M. and Gendler, S.J.
TITLE	Direct Submission
JOURNAL	Submitted (21-SEP-2001) Biochemistry, Mayo Clinic, 13400 E. Shea Boulevard, Scottsdale, AZ 85259, USA
FEATURES	<p>Location/Qualifiers</p> <p>1. .1834</p> <p>/organism="synthetic construct"</p> <p>/mol_type="mRNA"</p> <p>/db_xref="taxon:32630"</p> <p>/note="Homo sapiens gene in transgenic Mus musculus C57Bl/6; isolated from intestinal mucosa"</p> <p>1. .1834</p> <p>/gene="MUC1"</p> <p>1. .1452</p> <p>/gene="MUC1"</p> <p>/note="alternatively spliced; contains exon 6a resulting in variant hydrophilic tail of carboxy-terminal domain; lacks sites for beta-catenin and Gp2 interactions; derived from Homo sapiens"</p> <p>/codon_start=1</p> <p>/transl_table=1</p> <p>/product="mucin variant MUC-1CT80"</p> <p>/protein_id="AAL86734.1"</p> <p>/db_xref="GI:19338620"</p> <p>/translation="MTPGTQSPFFLLLLTVLTVTGSGHASTPGGEKETSATQRSSVPSSTEKNVSMSTSVLSHSPGSSSTGGQDVTLPATEPAGSAAWTGQDVTSPVTRPALGSTPPAHDVTSAPDNKPAGGSTAPPAHGVTSAADTRPAGSTAPPAHGVTSAPDNRPALGSTAPPAHNVTSASGSASASTLVHNGTSARATTTPASKSTPFSIPSHS SDTPTLASHSTKTDASTHSTVPLTSSNHSTSPQLSTGVSPFLSFHISNLOFNS SLEDPSTDYQELQRDISEMFLQYKQGGFLGLSNKFRPGSVVQVLTAFREGTINV HDVETQFNQYKTEASARNLTISDVSVDVPPFPFSAQSGAGVPGICALLVLCVLVA LAIVYLIALAVCCQRRKNYGLDIFPARDTYHPMSEYPTYHTGHVPPSSTRSPYE KVRLEPTGGQKQKRVNLGKDSGGTWTQRAWKR"</p>
ORIGIN	
Query Match	60.0%; Score 36; DB 12; Length 1834;
Best Local Similarity	75.0%; Pred. No. 0.44;
Matches	45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
Qy	1 GGTCAACGGCACCCTCCAGCACAGGAGTCACGTCGTGACCCGACACCCCGTCCAGTCCG 60
Db	385 GGCTCCACCGCCCCCCAGCCGCGGTGTCACTCGGCCCGGACACACGCGCGCCCCG 444
RESULT 13	
LOCUS	AX959918 1835 bp DNA linear PAT 14-JAN-2004
DEFINITION	Sequence 23 from Patent WO03100060.
ACCESSION	AX959918
VERSION	AX959918.1 GI:40880145
KEYWORDS	
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	Burden, N.G., Ellis, J.H. and Hamblin, P.A.
TITLE	Muc-1 antigen with reduced number of vnter repeat units
JOURNAL	Patent: WO 03100060-A 23 04-DEC-2003;
FEATURES	GLAXO GROUP LIMITED (GB)
source	Location/Qualifiers
	1. .1835
	/organism="Homo sapiens"
	/mol_type="unassigned DNA"
	/db_xref="taxon:9606"

Qy	1	GGTTCAACGGGACCTCCAGCACACGGAGTCACTGTGCACCCGACACCCCGTCCAGCTCCG	60	0.0%	Score 36;	DB 6;	Length 1835;
Db	462	GGCTCCACCGCCCCCCCCAGCCCAAGGTGTCACTCGCCCCGACACACGAGCCGCCCCG	521	75.0%	Pred. No. 0.44;		
Matches	45;	Conservative	0;	Mismatches	15;	Indels	0;
Gaps	0;						
RESULT 14							
LOCUS	AX959916						
DEFINITION	Sequence 21 from Patent WO03100060.						
ACCESSION	AX959916						
VERSION	AX959916.1						
KEYWORDS	GI:40880144						
SOURCE	Homo sapiens (human)						
ORGANISM	Homo sapiens						
REFERENCE	Burden,N.G., Ellis,J.H. and Hamblin,P.A.						
AUTHORS	Muc-1 antigen with reduced number of vnt repeat units						
TITLE	Patent: WO 03100060-A 21 04-DEC-2003;						
JOURNAL	GLAXO GROUP LIMITED (GB)						
FEATURES	Location/Qualifiers						
source	1. 2135						
	/organism="Homo sapiens"						
	/mol_type="unassigned DNA"						
	/db_xref="taxon:9606"						
ORIGIN							
Query Match	60.0%	Score 36;	DB 6;	Length 2135;			
Best Local Similarity	75.0%	Pred. No. 0.44;					
Matches	45;	Conservative	0;	Mismatches	15;	Indels	0;
Gaps	0;						
RESULT 15							
LOCUS	C0834017						
DEFINITION	Sequence 53 from Patent EPI439393.						
ACCESSION	C0834017						
VERSION	C0834017.1						
KEYWORDS	GI:50833622						
SOURCE	Homo sapiens (human)						
ORGANISM	Homo sapiens						
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.						
1	Asle,J.H., Boardman,L.A., Bugart,L.J., Burgess,C.C., Catino,T.J.,						
	Dwivedi,P., Huntress,M., Johnson,K.A., Lewis,M.E., Maimonis,P.J.,						
	Myerow,S.H., Brown-Shimer,S.L., Thiagalingam,A., Thibodeau,S.N. and						
	Molino,G.A.						
	Detection methods using TIMP 1 for colon cancer diagnosis						
	Patent: EP 1439393-A 53 21-JUL-2004;						
	Bayer Healthcare LLC (US); MAYO FOUNDATION FOR MEDICAL EDUCATION						
	AND RESEARCH (US)						
FEATURES	Location/Qualifiers						
source	1. 4139						
	/organism="Homo sapiens"						
	/mol_type="unassigned DNA"						
	/db_xref="taxon:9606"						
ORIGIN							
Query Match	60.0%	Score 36;	DB 6;	Length 4139;			
Best Local Similarity	75.0%	Pred. No. 0.41;					
Matches	45;	Conservative	0;	Mismatches	15;	Indels	0;
Gaps	0;						

QY 1 GGTCAACGGCACCCTCCAGCAGAGTCAAGTCTGACCCGACACCCGTCAGCTCCG 60
Db 458 GGCTCCACCGCCCCCAGCCCAAGGTGTCACTCGCCCGGACACAGGCGCGCCCG 517

Search completed: January 15, 2005, 22:36:54
Job time : 722.8 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 22:37:00 ; Search time 183.4 Seconds
(without alignments)
1879.789 Million cell updates/sec

Title: US-10-057-136-10
Perfect score: 60
Sequence: 1 gggctgactgccccctccggc.....ctgacacaagccagcccca 60

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	100.0	60	US-10-057-136-10	Sequence 10, Appl
2	60	100.0	2297	US-10-406-317-41	Sequence 41, Appl
3	38.2	63.7	60	US-10-057-136-2	Sequence 2, Appl
C	38.2	63.7	1424	US-10-447-839A-75	Sequence 75, Appl
4	38.2	63.7	1428	US-10-447-839A-20	Sequence 20, Appl
5	38.2	63.7	1527	US-10-057-136-19	Sequence 19, Appl
6	38.2	63.7	1527	US-10-057-136-19	Sequence 19, Appl
7	38.2	63.7	1799	US-10-447-839A-19	Sequence 19, Appl
8	38.2	63.7	1804	US-09-964-824A-573	Sequence 573, Appl
9	38.2	63.7	1804	US-10-029-517-17	Sequence 17, Appl
10	38.2	63.7	1804	US-10-717-597-30	Sequence 30, Appl
11	38.2	63.7	1804	US-10-775-920-84	Sequence 84, Appl
12	38.2	63.7	1823	US-10-101-510-339	Sequence 339, Appl

13	38.2	63.7	4139	9	US-09-964-824A-105	Sequence 105, App
14	38.2	63.7	4139	9	US-09-964-824A-578	Sequence 578, App
15	38.2	63.7	4139	9	US-09-864-864-334	Sequence 334, App
16	38.2	63.7	4139	9	US-09-880-107-2121	Sequence 2121, App
17	38.2	63.7	4139	11	US-09-968-007A-751	Sequence 751, App
18	38.2	63.7	4139	14	US-10-171-311-157	Sequence 157, App
19	38.2	63.7	4139	15	US-10-177-293-310	Sequence 310, App
20	38.2	63.7	4139	16	US-10-440-464-155	Sequence 155, App
21	38.2	63.7	4139	17	US-10-734-564-53	Sequence 53, App
22	38.2	63.7	4139	17	US-10-775-920-80	Sequence 80, App
23	38.2	63.7	4139	17	US-10-775-920-85	Sequence 85, App
24	38.2	63.7	8181	16	US-10-247-703-23	Sequence 23, App
25	38.2	63.7	8186	14	US-10-247-703-23	Sequence 23, App
26	38.2	63.7	8186	15	US-10-029-517-19	Sequence 19, App
27	36.6	61.0	572	15	US-10-029-517-18	Sequence 18, App
28	36.6	61.0	572	17	US-10-775-920-93	Sequence 93, App
29	36.6	61.0	1721	9	US-09-864-864-280	Sequence 280, App
30	36.6	61.0	1721	9	US-09-967-768A-224	Sequence 224, App
31	36.6	61.0	1721	14	US-10-247-703-21	Sequence 21, App
32	36.6	61.0	1721	14	US-10-097-340-211	Sequence 211, App
33	36.6	61.0	1721	14	US-10-171-311-155	Sequence 155, App
34	36.6	61.0	1721	15	US-10-007-926A-58	Sequence 58, App
35	36.6	61.0	1721	15	US-10-029-517-3	Sequence 3, Appl
36	36.6	61.0	1721	15	US-10-172-118-775	Sequence 775, App
37	36.6	61.0	1721	16	US-10-342-887-775	Sequence 775, App
38	36.6	61.0	1721	17	US-10-775-920-88	Sequence 88, Appl
39	36.6	61.0	2026	14	US-10-198-846-12589	Sequence 12589, A
40	36.6	61.0	2238	17	US-10-775-920-87	Sequence 87, Appl
41	36.6	61.0	2678	15	US-10-252-157-103	Sequence 103, Appl
42	36.2	60.3	60	14	US-10-057-136-14	Sequence 14, Appl
43	36.2	60.3	518	14	US-10-247-703-38	Sequence 38, Appl
44	36.2	60.3	518	15	US-10-029-517-101	Sequence 101, App
45	36.2	60.3	981	15	US-10-029-517-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-10-057-136-10
; Sequence 10, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUPF, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: TUMOR-ASSOCIATED ANTIGEN
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-10

Query Match 100.0%; Score 60; DB 14; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.1e-12;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCAGCTCTGACACAGGCGACCCCA 60
|||||

Db 1 GGGTCGACTGCCCTCCGGCGCATGTGTGACCTCAGCTCCTGACACAAAGCCAGCCCCA 60

RESULT 2

US-10-406-317-41
; Sequence 41, Application US/10406317
; Publication No. US20040019195A1
; GENERAL INFORMATION:
; APPLICANT: Schlom, Jeffrey;
; APPLICANT: Hodge, James;
; APPLICANT: Panicali, Dennis
; TITLE OF INVENTION: A recombinant vector expressing multiple constitutulatory
; FILE REFERENCE: 38163-0189
; CURRENT APPLICATION NUMBER: US/10/406,317
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US/09/856,988
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: PCT/US99/26866
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/111,582
; PRIOR FILING DATE: 1998-12-09
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: VECTOR
US-10-406-317-41

Query Match 100.0%; Score 60; DB 16; Length 2297;
Best Local Similarity 100.0%; Pred. No. 1.1e-12;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTCGACTGCCCTCCGGCGCATGTGTGACCTCAGCTCCTGACACAAAGCCAGCCCCA 60
Db 766 GGGTCGACTGCCCTCCGGCGCATGTGTGACCTCAGCTCCTGACACAAAGCCAGCCCCA 825

RESULT 3

US-10-057-136-2
; Sequence 2, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-2

Query Match 63.7%; Score 38.2; DB 14; Length 60;
Best Local Similarity 78.0%; Pred. No. 0.00014;

Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGGTCGACTGCCCTCCGGCGCATGTGTGACCTCAGCTCCTGACACAAAGCCAGCCCC 59
Db 1 GGCTCCACCGCCCCCGCCAGCCACGGTGTACCTCGGCCCGGACACCAAGCGCGCCCC 59

RESULT 4

US-10-447-839A-75/c
; Sequence 75, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:
; APPLICANT: KuFe, Donald W.
; APPLICANT: Kharbanda, Surender
; APPLICANT: Weltman, Steven D.
; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT APPLICATION NUMBER: US/10/447,839A
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 75
; LENGTH: 1424
; TYPE: RNA
; ORGANISM: ARTIFICIAL
; FEATURE:
; OTHER INFORMATION: Synthesized Sequence
US-10-447-839A-75

Query Match 63.7%; Score 38.2; DB 16; Length 1424;
Best Local Similarity 78.0%; Pred. No. 0.00015;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGGTCGACTGCCCTCCGGCGCATGTGTGACCTCAGCTCCTGACACAAAGCCAGCCCC 59
Db 1041 GGCTCCACCGCCCCCGCCAGCCACGGTGTACCTCGGCCCGGACACCAAGCGCGCCCC 983

RESULT 5

US-10-447-839A-20
; Sequence 20, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:
; APPLICANT: KuFe, Donald W.
; APPLICANT: Kharbanda, Surender
; APPLICANT: Weltman, Steven D.
; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT APPLICATION NUMBER: US/10/447,839A
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 1428
; TYPE: RNA
; ORGANISM: RNA
US-10-447-839A-20

Query Match 63.7%; Score 38.2; DB 16; Length 1428;
Best Local Similarity 71.2%; Pred. No. 0.00015;
Matches 42; Conservative 4; Mismatches 13; Indels 0; Gaps 0;


```
; Publication No. US20040110221A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael E.
; APPLICANT: Twine, Natalie C.
; APPLICANT: Dörner, Andrew J.
; APPLICANT: Trepicchio, William L.
; APPLICANT: Stonim, Donna K.
; APPLICANT: Stever, Jennifer A.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING RCC AND OTHER SOLID TUMORS
; FILE REFERENCE: AM101080L
; CURRENT APPLICATION NUMBER: US/10/717,597
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US 60/459,782
; PRIOR FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: US 60/427,982
; PRIOR FILING DATE: 2002-11-21
; NUMBER OF SEQ ID NOS: 4904
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-717-597-30
```

```
Query Match
Best Local Similarity 63.7%; Score 38.2; DB 17; Length 1804;
Matches 46; Conservativity 0; Mismatches 13; Indels 0; Gaps 0;
```

```
OY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCCTCTGACACACAGGCCAGCCCC 59
DB 457 GGCTCCACCGCCCCCGCCAGCCACGGGTGTCACTTCGGCCCCGGACACACAGGCCAGCCCC 515
```

```
RESULT 11
US-10-775-920-84
; Sequence 84, Application US/10775920
; Publication No. US20040175744A1
; GENERAL INFORMATION:
; APPLICANT: Merigen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; FILE REFERENCE: Merigen - 0010B
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 84
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-775-920-84
```

```
Query Match
Best Local Similarity 63.7%; Score 38.2; DB 17; Length 1804;
Matches 46; Conservativity 78.0%; Pred. No. 0.00015;
Mismatches 0; Mismatches 13; Indels 0; Gaps 0;
```

```
OY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCCTCTGACACACAGGCCAGCCCC 59
DB 457 GGCTCCACCGCCCCCGCCAGCCACGGGTGTCACTTCGGCCCCGGACACACAGGCCAGCCCC 515
```

```
RESULT 12
US-10-101-510-339
; Sequence 339, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
```

```
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 339
; LENGTH: 1823
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-101-510-339
```

```
Query Match
Best Local Similarity 63.7%; Score 38.2; DB 15; Length 1823;
Matches 46; Conservativity 78.0%; Pred. No. 0.00015;
Mismatches 0; Mismatches 13; Indels 0; Gaps 0;
```

```
OY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCCTCTGACACAGGCCAGCCCC 59
DB 457 GGCTCCACCGCCCCCGCCAGCCACGGGTGTCACTTCGGCCCCGGACACACAGGCCAGCCCC 515
```

```
RESULT 13
US-09-964-824A-105
; Sequence 105, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horrigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signat
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964,824A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 105
; LENGTH: 4139
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-964-824A-105
```

```
Query Match
Best Local Similarity 63.7%; Score 38.2; DB 9; Length 4139;
Matches 46; Conservativity 78.0%; Pred. No. 0.00015;
Mismatches 0; Mismatches 13; Indels 0; Gaps 0;
```

```
OY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCCTCTGACACAGGCCAGCCCC 59
DB 458 GGCTCCACCGCCCCCGCCAGCCACGGGTGTCACTTCGGCCCCGGACACACAGGCCAGCCCC 516
```

```
RESULT 14
US-09-964-824A-578
; Sequence 578, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horrigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964,824A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
; PRIOR FILING DATE: 2000-09-28
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 19:50:13 : Search time 35.2 Seconds
(without alignment)
1211.572 Million cell updates/sec

Title: US-10-057-136-10

Perfect score: 60
Sequence: 1 GGGTCGACTGCCCTCCGCGC.....CTGACACAAGCCAGCCCA 60

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38.2	63.7	1804	US-10-029-517-17	Sequence 17, Appl
2	38.2	63.7	8186	US-10-029-517-19	Sequence 19, Appl
3	36.6	61.0	572	US-10-029-517-18	Sequence 18, Appl
4	36.6	61.0	1721	US-10-029-517-3	Sequence 3, Appl
5	36.2	60.3	518	US-10-029-517-101	Sequence 101, App
6	36.2	60.3	981	US-10-029-517-16	Sequence 16, Appl
7	36.2	60.3	3343	US-10-029-517-102	Sequence 102, App
8	35	58.3	6192	US-08-479-537A-1	Sequence 1, Appl
9	35	58.3	6192	US-09-083-116-1	Sequence 1, Appl
10	35	58.3	6192	US-09-134-916A-1	Sequence 1, Appl
11	35	58.3	6449	US-08-479-537A-4	Sequence 4, Appl
12	35	58.3	6449	US-09-083-116-4	Sequence 4, Appl
13	35	58.3	6449	US-09-134-916A-4	Sequence 4, Appl
14	33.8	56.3	60	US-09-475-947A-246	Sequence 246, App
15	24.8	41.3	4176	US-09-722-139-1	Sequence 1, Appl
16	24.8	41.3	4176	US-09-721-832-1	Sequence 1, Appl
17	24.8	41.3	4176	US-09-721-689-1	Sequence 1, Appl
18	24.8	41.3	4767	US-09-231-899-76	Sequence 76, Appl
19	24.4	40.7	467	US-09-016-434-1158	Sequence 1158, Ap
20	24.4	40.7	556	US-10-029-517-105	Sequence 105, App
21	24.4	40.7	564	US-09-621-976-3738	Sequence 3738, Ap
22	24.4	40.7	1155	US-09-016-434-1455	Sequence 1455, Ap
23	24	40.0	4473	US-08-894-173-1	Sequence 1, Appl
24	24	40.0	4473	US-09-398-193-1	Sequence 1, Appl
25	23.8	39.7	2769	US-09-489-039A-1767	Sequence 1767, Ap
26	23.6	39.3	519	US-09-646-028-42	Sequence 42, Appl
27	23.6	39.3	534	US-09-646-028-46	Sequence 46, Appl

28	23.4	39.0	346	4	US-09-513-999C-251	Sequence 251, App
29	23.4	39.0	416	4	US-09-513-999C-8438	Sequence 8438, Ap
C	30	23.4	1163	4	US-09-522-714-5	Sequence 5, Appl
	31	23.4	39.0	6	5340934-1	Patent No. 5340934
C	32	23.2	258	4	US-09-513-999C-9477	Sequence 9477, Ap
C	33	23.2	663	3	PCT-US93-06251-21	Sequence 1067, Ap
C	34	23.2	695	5	US-08-306-691B-14	Sequence 21, Appl
C	35	23.2	6453	1	US-09-209-668-10	Sequence 10, Appl
C	36	23.2	6453	3	US-09-266-965-76	Sequence 8, Appl
C	37	23.2	6453	3	US-09-356-952-8	Sequence 76, Appl
C	38	23.2	53500	4	US-09-270-767-14564	Sequence 14564, A
C	39	23	1240	4	US-08-553-110-2	Sequence 2, Appl
	40	23	1853	1	US-09-266-225D-17	Sequence 17, Appl
	41	23	4168	4	US-09-041-886-22	Sequence 22, Appl
	42	23	4279	3	US-09-621-976-8387	Sequence 8387, Ap
C	43	22.8	401	4	US-09-336-536-16	Sequence 16, Appl
C	44	22.8	1218	4	US-09-336-536-15	Sequence 15, Appl
C	45	22.8	1831	4	US-09-336-536-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-10-029-517-17
Sequence 17, Application US/10029517
Patent No. 6716627
GENERAL INFORMATION:
APPLICANT: Kenneth W. Dobie
APPLICANT: Susan J. Myers
TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
FILE REFERENCE: RTS-0352
CURRENT APPLICATION NUMBER: US/10/029,517
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 107
SEQ ID NO 17
LENGTH: 1804
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (73)...(1500)
US-10-029-517-17

Query Match 63.7%; Score 38.2; DB 4; Length 1804;
Best Local Similarity 78.0%; Pred. No. 0.0011;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGGTCGACTGCCCTCCGCGCATGCTGACCTCCTGACACAAGCCAGCCCC 59
Db 457 GGCTCACCGCCCGCCCGACGCGTGTACCTCGCGCCCGACACAGCGGCCCC 515

RESULT 2
US-10-029-517-19
Sequence 19, Application US/10029517
Patent No. 6716627
GENERAL INFORMATION:
APPLICANT: Kenneth W. Dobie
APPLICANT: Susan J. Myers
TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
FILE REFERENCE: RTS-0352
CURRENT APPLICATION NUMBER: US/10/029,517
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 107
SEQ ID NO 19
LENGTH: 8186
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: unsure
LOCATION: 6899
OTHER INFORMATION: unknown

```
/ NAME/KEY: unsure
/ LOCATION: 7155
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 7184
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 7957
/ OTHER INFORMATION: unknown
/ NAME/KEY: intron
/ LOCATION: (2997)...(3498)
/ OTHER INFORMATION: intron 1
/ NAME/KEY: intron:exon junction
/ LOCATION: (3498)...(3499)
/ OTHER INFORMATION: intron 1:exon 2
/ NAME/KEY: exon
/ LOCATION: (3508)...(3599)
/ OTHER INFORMATION: exon 2d
/ NAME/KEY: exon:intron junction
/ LOCATION: (3982)...(3983)
/ OTHER INFORMATION: exon 2a:intron 2a
/ NAME/KEY: intron:exon junction
/ LOCATION: (4205)...(4206)
/ OTHER INFORMATION: intron 2c:exon 3c
/ NAME/KEY: intron:exon junction
/ LOCATION: (4259)...(4260)
/ OTHER INFORMATION: intron 2d:exon 3d
/ NAME/KEY: exon
/ LOCATION: (4260)...(4328)
/ OTHER INFORMATION: exon 3d
/ NAME/KEY: intron:exon junction
/ LOCATION: (4632)...(4633)
/ OTHER INFORMATION: intron 3:exon 4
/ NAME/KEY: exon
/ LOCATION: (4914)...(5035)
/ OTHER INFORMATION: exon 5
/ NAME/KEY: intron
/ LOCATION: (5266)...(6293)
/ OTHER INFORMATION: intron 6
US-10-029-517-19
```

Query Match 63.7%; Score 38.2; DB 4; Length 8186;
Best Local Similarity 78.0%; Pred. No. 0.0014;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```
OY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCAGCTCCTGACACAAGGCCAGCCCC 59
Db 3825 GGCTCCACCGCCCCCCCCCAGCCCAAGGTTGACCTCGGCCCCGAGACCAAGCGCGCCCC 3883
```

```
RESULT 3
US-10-029-517-18
/ Sequence 18, Application US/10029517
/ Patent No. 6716627
/ GENERAL INFORMATION:
/ APPLICANT: Kenneth W. Dobie
/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
/ FILE REFERENCE: RTS-0352
/ CURRENT APPLICATION NUMBER: US/10/029,517
/ NUMBER OF SEQ ID NOS: 107
/ SEQ ID NO 18
/ LENGTH: 572
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (67)...(572)
US-10-029-517-18
```

Query Match 61.0%; Score 36.6; DB 4; Length 572;
Best Local Similarity 76.3%; Pred. No. 0.0033;

```
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
OY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCAGCTCCTGACACAAGGCCAGCCCC 59
Db 478 GGCTCCACCGCCCCCAAGCCCAAGCGTGTACCTCGGCCCCGAGACCAAGCGCGCCCC 536
```

```
RESULT 4
US-10-029-517-3
/ Sequence 3, Application US/10029517
/ Patent No. 6716627
/ GENERAL INFORMATION:
/ APPLICANT: Susan J. Myers
/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
/ FILE REFERENCE: RTS-0352
/ CURRENT APPLICATION NUMBER: US/10/029,517
/ NUMBER OF SEQ ID NOS: 107
/ SEQ ID NO 3
/ LENGTH: 1721
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (58)...(1605)
US-10-029-517-3
```

Query Match 61.0%; Score 36.6; DB 4; Length 1721;
Best Local Similarity 76.3%; Pred. No. 0.0038;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

```
OY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCAGCTCCTGACACAAGGCCAGCCCC 59
Db 442 GGCTCCACCGCCCCCCCCCAGCCCAAGCGTGTACCTCGGCCCCGAGACCAAGCGCGCCCC 500
```

```
RESULT 5
US-10-029-517-101
/ Sequence 101, Application US/10029517
/ Patent No. 6716627
/ GENERAL INFORMATION:
/ APPLICANT: Kenneth W. Dobie
/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
/ FILE REFERENCE: RTS-0352
/ CURRENT APPLICATION NUMBER: US/10/029,517
/ NUMBER OF SEQ ID NOS: 107
/ SEQ ID NO 101
/ LENGTH: 518
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
US-10-029-517-101
```

Query Match 60.3%; Score 36.2; DB 4; Length 518;
Best Local Similarity 77.2%; Pred. No. 0.0044;
Matches 44; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```
OY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCAGCTCCTGACACAAGGCCAGGCC 57
Db 462 GGCTCCACCGCCCCCAAGCCCAAGCGTGTACCTCGGCCCCGAGACCAAGCGCGGCC 518
```

```
RESULT 6
US-10-029-517-16
/ Sequence 16, Application US/10029517
/ Patent No. 6716627
/ GENERAL INFORMATION:
/ APPLICANT: Kenneth W. Dobie
/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
```

```
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 16
; LENGTH: 981
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon:exon junction
; LOCATION: (464)..(465)
; OTHER INFORMATION: exon 3b:exon 4
US-10-029-517-16

Query Match          60.3%; Score 36.2; DB 4; Length 981;
Best Local Similarity 77.2%; Pred. No. 0.0047;
Matches 44; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCCTGACACAGGCCAGCC 57
    |||||
Db 21 GGCTCCACCGCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 77

RESULT 7
US-10-029-517-102
; Sequence 102, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 102
; LENGTH: 3343
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-10-029-517-102

Query Match          60.3%; Score 36.2; DB 4; Length 3343;
Best Local Similarity 77.2%; Pred. No. 0.0055;
Matches 44; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCCTGACACAGGCCAGCC 57
    |||||
Db 1728 GGCTCCACCGCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 1784

RESULT 8
US-08-479-537A-1
; Sequence 1, Application US/08479537A
; Patent No. 5861381
; GENERAL INFORMATION:
; APPLICANT: CHAMON, Pierre
; APPLICANT: KIENY, Marie-Paule
; APPLICANT: LATHE, Richard
; APPLICANT: HAREUVENT, Maira
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
; TITLE OF INVENTION: TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,537A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 90/13101
; FILING DATE: 23-OCT-1990
; APPLICATION NUMBER: WO PCT/FR91/00835
; FILING DATE: 23-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/039,320
; FILING DATE: 04-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,576
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 017753-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6192 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 58..120
; NAME/KEY: repeat_region
; LOCATION: 439..5239
; OTHER INFORMATION: /note= "The nucleotides spanning
; OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is
; OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixe
; OTHER INFORMATION: The number of such repeats varies from 1 to 80."
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 121..6166
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 457
; OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
; OTHER INFORMATION: which is the codon for Pro or Asn wherein Thr = ACT, ACC, ACY
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 487
; OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
; OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACY
; OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
; NAME/KEY: repeat_region
; LOCATION: 496
; OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCP
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-08-479-537A-1

Query Match          58.3%; Score 35; DB 2; Length 6192;
Best Local Similarity 74.6%; Pred. No. 0.015;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCCTGACACAGGCCAGCC 59
    |||||
Db 382 GGCTCCACCGCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 440
```

```
RESULT 9
US-09-083-116-1
; Sequence 1, Application US/09083116
; Patent No. 6203795
; GENERAL INFORMATION:
; APPLICANT: CHAMON, Pierre
; APPLICANT: KIENY, Marie-Paule
; APPLICANT: LATHE, Richard
; APPLICANT: HAREUVENI, Mara
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
; TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/083,116
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,537
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR91/00835
; FILING DATE: 23-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/039,320
; FILING DATE: 04-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,576
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 017753-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6192 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 58..120
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 439..5239
; OTHER INFORMATION: /note= "The nucleotides spanning
; 439-5239 constitute a repeated region wherein the repeat is 6
; OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
; OTHER INFORMATION: The number of such repeats varies from 1 to 80."
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 121..6166
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 457
; OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
```

```
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 487
; OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
; OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
; OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 496
; OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; US-09-083-116-1
;
Query Match      58.3%; Score 35; DB 3; Length 6192;
Best Local Similarity 74.6%; Pred. No. 0.015;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
;
QY      1  GGGTCGACTGCCCTCCGGCGCATGTGTGACCTCAGCTCTGACACAGGCCAGCCCC 59
Db      382 GGCTCCACCACCCCGCCAGCCACGATGTCACTCAGCCCGGACACCAAGCCAGCCCC 440
;
RESULT 10
US-09-134-916A-1
; Sequence 1, Application US/09134916A
; Patent No. 6328956
; GENERAL INFORMATION:
; APPLICANT: CHAMON, Pierre
; APPLICANT: KIENY, Marie-Paule
; APPLICANT: LATHE, Richard
; APPLICANT: HAREUVENI, Mara
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
; TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/134,916A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,537
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: FR 90/13101
; FILING DATE: 23-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR91/00835
; FILING DATE: 23-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/039,320
; FILING DATE: 04-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,576
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 017753-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
```

```

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6192 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 58..120
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 439..5239
; OTHER INFORMATION: /note= "The nucleotides spanning
; OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
; OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
; OTHER INFORMATION: The number of such repeats varies from 1 to 80."
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 121..6166
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 457
; OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 487
; OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
; OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
; OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 496
; OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; US-09-134-916A-1
;
; Query Match 58.3%; Score 35; DB 3; Length 6192;
; Best Local Similarity 74.6%; Pred. No. 0.015;
; Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
;
Qy 1 GGGTCGACTGCCCTCCGGCGCATGTGTGACCTCAGCTCCTGACACAAGCGCAGCCCC 59
Db 382 GGCTCCACCAACCCCGCCGACGATGTCACTCAGCCCGGACACAAGCCAGCCCC 440
;
; RESULT 11
; US-08-479-537A-4
; Sequence 4, Application US/08479537A
; Patent No. 5861381
; GENERAL INFORMATION:
; APPLICANT: CHAMON, Pierre
; APPLICANT: KIENY, Marie-Paule
; APPLICANT: LATHE, Richard
; APPLICANT: HAREUVENI, Mera
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
; TITL OF INVENTION: TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479, 537A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 90/13101
; FILING DATE: 23-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR91/00835
; FILING DATE: 23-OCT-1991
; APPLICATION NUMBER: US 08/039,320
; FILING DATE: 04-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,576
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 017753-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6449 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 58..120
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 439..5239
; OTHER INFORMATION: /note= "The nucleotides spanning
; OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
; OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
; OTHER INFORMATION: The number of such repeats varies from 1 to 80."
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 121..5661
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 457
; OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 487
; OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
; OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
; OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 496
; OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; US-08-479-537A-4
;
; Query Match 58.3%; Score 35; DB 2; Length 6449;
; Best Local Similarity 74.6%; Pred. No. 0.015;
; Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
;
Qy 1 GGGTCGACTGCCCTCCGGCGCATGTGTGACCTCAGCTCCTGACACAAGCGCAGCCCC 59
Db 382 GGCTCCACCAACCCCGCCGACGATGTCACTCAGCCCGGACACAAGCCAGCCCC 440
;
; RESULT 12

```


US-09-083-116-4
; Sequence 4, Application US/09083116
; Patent No. 6203795
; GENERAL INFORMATION:
; APPLICANT: CHAMON, Pierre
; APPLICANT: KIENY, Marie-Paule
; APPLICANT: LATHE, Richard
; APPLICANT: HAREUVENI, Mara
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
; TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/083,116
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,537
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR91/00835
; FILING DATE: 23-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/039,320
; FILING DATE: 04-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,576
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 017753-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6449 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 58..120
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 439..5239
; OTHER INFORMATION: /note= "The nucleotides spanning
; 439-5239 constitute a repeated region wherein the repeat is 6
; nucleotides and encodes 20 amino acids, 17 of which are fixed
; OTHER INFORMATION: The number of such repeats varies from 1 to 80."
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 121..5661
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 457
; OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; FEATURE:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:

NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-09-083-116-4
Query Match 58.3%; Score 35; DB 3; Length 6449;
Best Local Similarity 74.6%; Pred. No. 0.015;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 GGCTCGACTGCCCTCCGGCGATGTGTACCTCAGCTTCTGACACAGGCCAGCCCC 59
Db 382 GGCTCCACGACCCCGCCAGCCGACGATGTCACTCAGCCCCGACACAGGCCAGCCCC 440
RESULT 13
US-09-134-916A-4
; Sequence 4, Application US/09134916A
; Patent No. 6328956
; GENERAL INFORMATION:
; APPLICANT: CHAMON, Pierre
; APPLICANT: KIENY, Marie-Paule
; APPLICANT: LATHE, Richard
; APPLICANT: HAREUVENI, Mara
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
; TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/134,916A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,537
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: FR 90/13101
; FILING DATE: 23-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR91/00835
; FILING DATE: 23-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/039,320
; FILING DATE: 04-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,576
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 017753-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 19:43:55 : Search time 1550.6 Seconds
(without alignments)
1410.024 Million cell updates/sec

Title: US-10-057-136-10

Perfect score: 60
Sequence: 1 gggtcgactgcctccctccggc.....ctgacacaagcgccaccacca 60

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	39.8	66.3	330	1	AI925867 wo20d04.x
2	39.8	66.3	1262	5	BQ935496 AGENCOURT
3	39.8	66.3	1349	5	BU152566 AGENCOURT
4	38.2	63.7	604	4	BM791359
5	38.2	63.7	877	5	BU542454
6	38.2	63.7	959	6	CA489836
7	38.2	63.7	1113	5	BU148487
8	38.2	63.7	1130	5	BU542996
9	38.2	63.7	1234	5	BQ936898
10	38.2	63.7	1268	5	BQ943554
11	38.2	63.7	1343	5	BQ920055
12	38.2	63.7	1420	5	BU542790
13	38.2	63.7	1531	5	BU543309
14	36.2	60.3	981	4	BG774910
15	36.2	60.3	1536	5	BQ923149
16	35.2	58.7	166	7	T27692
17	35.2	58.7	475	6	CB120860
18	35.2	58.7	669	6	CB122585
19	35	58.3	472	4	BM759495
20	32	53.3	1334	4	BQ943809
21	30.2	50.3	1678	4	BG775565
22	29.8	49.7	308	7	R48730
23	29.6	49.3	619	4	BI260921
24	29.2	48.7	548	2	BE706360

25	27.8	46.3	528	4	BM130288	BM130288 pb29b06.y
26	27.8	46.3	535	4	BI360713	BI360713 388731 MA
27	27.8	46.3	538	4	BI744470	BI744470 pb09c12.y
28	27.2	45.3	270	5	BY518871	BY518871 BY518871
29	27	45.0	616	6	CF176642	CF176642 805040 MA
30	26.8	44.7	300	8	AO310138	AO310138 CITBI-B1-
31	26.8	44.7	398	8	B43272	B43272 HS-1057-A2-
c 32	26.8	44.7	401	2	BE327251	BE327251 hw09d04.x
33	26.8	44.7	441	8	B43540	B43540 HS-1057-B2-
c 34	26.4	44.0	617	6	CD218221	CD218221 pgr1n.pk0
c 35	26.4	44.0	617	7	CN222357	CN222357 WLA040D03
36	26.4	44.0	838	9	CG306143	CG306143 OGMHY94TV
37	26.4	44.0	847	3	AY109264	AY109264 Zea mays
38	26.4	44.0	898	9	CG053439	CG053439 PUKAY07TB
39	26.4	44.0	937	7	CG249534	CG249534 zmtw005
40	26.4	44.0	982	9	CG249534	CG249534 OGI8C68TV
c 41	26.4	44.0	1172	7	CK209164	CK209164 FGAS02091
42	26.2	43.7	298	6	CA803305	CA803305 ESG011C.E
43	26.2	43.7	630	5	BQ199203	BQ199203 UI-R-DQ1-
c 44	26.2	43.7	993	6	CA082679	CA082679 SCBFAM202
45	26	43.3	358	5	BY466019	BY466019 BY466019

ALIGNMENTS

RESULT 1
AI925867/c 330 bp mRNA linear EST 08-MAR-2000
LOCUS wo20d04.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2455879 3'
DEFINITION similar to contains element MSRI MSRI repetitive element ;, mRNA
sequence.

ACCESSION AI925867
VERSION AI925867
KEYWORDS AI925867.1 GI:5661831
SOURCE EST.

ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 330)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov

Life Technologies catalog #: 11548-013

DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www-bio.lnl.gov/bdrp/image/image.html

Insert length: 578 Std Error: 0.00
Seq primer: -40UP from Gibco

High quality sequence stop: 305.
Location/Qualifiers

FEATURES

source

1..330
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2455879"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Pan1"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.72 kb. Life Technologies catalog #:
11548-013"

ORIGIN

Query Match 66.3%; Score 39.8; DB 1; Length 330;
Best Local Similarity 79.7%; Pred. No. 0.028;
Matches 47; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 GGGTCGACTGCCCTCCGGCGCATGTGTGACCTCAGCTCCTGACACAGAGCCAGCCCC 59
|||
Db 256 GGCTCCACCGCCCCCCCCCAGCCGAGGTGTACCTCGGCTCCGGACACACAGCGCGCCCC 198
|||
RESULT 2
BQ935496 1262 bp mRNA linear EST 21-AUG-2002
LOCUS AGENCOURT_8776475 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6382088
DEFINITION 5', mRNA sequence.
ACCESSION BQ935496
VERSION BQ935496.1 GI:22350879
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1262)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLCM2573 row: f column: 09
High quality sequence start: 46
High quality sequence stop: 157.
Location/Qualifiers
1..1262
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6382088"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 66.3%; Score 39.8; DB 5; Length 1262;
Best Local Similarity 79.7%; Pred. No. 0.029;
Matches 47; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 GGGTCGACTGCCCTCCGGCGCATGTGTGACCTCAGCTCCTGACACAGAGCCAGCCCC 59
|||
Db 56 GGCTCCACCGCCCCCCCCCAGCCGAGGTGTACCTCGGCTCCGGACACACAGCGCGCCCC 114
|||

RESULT 3
BUI52566 1349 bp mRNA linear EST 03-SEP-2002
LOCUS AGENCOURT_8670683 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6380559
DEFINITION 5', mRNA sequence.
ACCESSION BUI52566
VERSION BUI52566.1 GI:22666098
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 1349)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLCM2569 row: f column: 16
High quality sequence stop: 291.
Location/Qualifiers
1..1349
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6380559"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 66.3%; Score 39.8; DB 5; Length 1349;
Best Local Similarity 79.7%; Pred. No. 0.029;
Matches 47; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 GGGTCGACTGCCCTCCGGCGCATGTGTGACCTCAGCTCCTGACACAGAGCCAGCCCC 59
|||
Db 259 GGCTCCACCGCCCCCCCCCAGCCGAGGTGTACCTCGGCTCCGGACACACAGCGCGCCCC 317
|||

RESULT 4
BM791359 604 bp mRNA linear EST 05-MAR-2002
LOCUS K-EST0071342 S21SNU520 Homo sapiens cDNA clone S21SNU520-14-A06 5',
DEFINITION mRNA sequence.
ACCESSION BM791359
VERSION BM791359.1 GI:19139591
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 604)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kr.ibm.re.kr
Plate: 14 row: A column: 06
High quality sequence stop: 604.
Location/Qualifiers
1..604
/organism="Homo sapiens"

RESULT 7
BU148487
LOCUS
DEFINITION BU148487 1113 bp mRNA linear EST 03-SEP-2002
AGENCOURT 8670479 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6380642
5', mRNA Sequence.
ACCESSION BU148487
VERSION
KEYWORDS BU148487.1 GI:22662019
EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE NIH-MGC <http://mgs.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM2569 row: j column: 03
High quality sequence stop: 235.
Location/Qualifiers
1..1113
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6380642"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

Query Match 63.7%; Score 38.2; DB 5; Length 1113;
Best Local Similarity 78.0%; Pred. No. 0.092;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGGTCGACTGCCCTCCGGCGCATGTGTGACCTCAGCTCCTGACACAGGCCAGCCCC 59
Db 30 GGCTCCACCGCCCCCGCCAGCCACGGTGTCACTCTGCCCGGACACCAAGCGCGCCCC 88

RESULT 8
BU542996
LOCUS
DEFINITION BU542996 1130 bp mRNA linear EST 13-SEP-2002
AGENCOURT_10338707 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574979
5', mRNA sequence.
ACCESSION BU542996
VERSION
KEYWORDS BU542996.1 GI:22853479
EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE NIH-MGC <http://mgs.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM2771 row: c column: 11
High quality sequence start: 27
High quality sequence stop: 246.
Location/Qualifiers
1..1130
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6574979"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

Query Match 63.7%; Score 38.2; DB 5; Length 1130;
Best Local Similarity 78.0%; Pred. No. 0.092;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGGTCGACTGCCCTCCGGCGCATGTGTGACCTCAGCTCCTGACACAGGCCAGCCCC 59
Db 69 GGCTCCACCGCCCCCGCCAGCCACGGTGTCACTCTGCCCGGACACCAAGCGCGCCCC 127

RESULT 9
BQ936898
LOCUS
DEFINITION BQ936898 1234 bp mRNA linear EST 21-AUG-2002
AGENCOURT_8919110 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6383855
5', mRNA sequence.
ACCESSION BQ936898
VERSION
KEYWORDS BQ936898.1 GI:22352281
EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE NIH-MGC <http://mgs.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM2577 row: o column: 24
High quality sequence stop: 245.
Location/Qualifiers
1..1234
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6383855"

ORIGIN

Qy 1 GGGTCGACTGCCCTCCGGCGCATGGTGTGACCTCAGCTCCTGACACAAGGCCAGCCCC 59
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 30 GGCTCCACCAGCCCCCAAGCCCAAGGTGTCAACCTCTGCACCGAACAAGGCCGCCCC 88

LOCUS	B0943554	1268 bp	mRNA	linear	EST 21-AUG-2002
DEFINITION	AGENCOURT_8777815 NIH_MGC_40 Homo sapiens CDNA clone IMAGE:6384308 5', mRNA Sequence.				

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

FEATURES

ORIGIN

qy 1 GGTGACTGCCCCCTCCGGGCATGGTGTGACCTCAGCTCTGACACAGGCCAGCCCC 59

Db

90 GGCCTCACGCCCCCCCCCAGCCCAAGGTCTCACTCGGCCCCGACACCAAGGCGGCCCC 148

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

JOURNAL COMMENT
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

FEATURES

ORIGIN

Best Local Similarity 78.0%; Pred. No. 0.093;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Db 72 GGCTCCACGCCCCCCCCAGCCACGGTGTCACTCTGCCCCGGACACACGAGCCGGCCCC 130

DEFINITION AGENCOURT_10334841 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574725 5', mRNA sequence.

ACCESSION	BU542790
VERSION	BU542790.1
	GI:22853273

REFERENCE
1 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1420)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
<http://image.llnl.gov>
Plate: L1CM2770 row: h column: 21
High quality sequence stop: 288.
Location/Qualifiers

FEATURES
source 1..1420
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6574725"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 63.7%; Score 38.2; DB 5; Length 1420;
Best Local Similarity 78.0%; Pred. No. 0.093;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 GGGTCGACTGCCCCCTCCGGCGCATGCTGACCTCAGCTCTGACACAGGCCAGCCCC 59
|||
Db 19 GGCTCCACCGCCCCCGCCAGCCCGGTGCTCAGCTCGGCCCGGACACAGGCCAGCCCC 77

RESULT 13
BUS43309 1531 bp mRNA linear EST 13-SEP-2002
LOCUS AGENCOURT_10327072 NIH_MGC_40 Homo sapiens CDNA clone IMAGE:6575349
ACCESSION BUS43309
VERSION BUS43309
KEYWORDS BUS43309.1 GI:22853792
SOURCE EST.
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NIH-MGC <http://mgc.nci.nih.gov/>.
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
<http://image.llnl.gov>
Plate: L1CM2772 row: b column: 21
High quality sequence stop: 166.
Location/Qualifiers

FEATURES

source 1..1531
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"
/clone="IMAGE:6575349"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 63.7%; Score 38.2; DB 5; Length 1531;
Best Local Similarity 78.0%; Pred. No. 0.093;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 GGGTCGACTGCCCCCTCCGGCGCATGCTGACCTCAGCTCTGACACAGGCCAGCCCC 59
|||
Db 30 GGCTCCACCGCCCCCGCCAGCCCGGTGCTCAGCTCGGCCCGGACACAGGCCAGCCCC 88

RESULT 14
BG774910 981 bp mRNA linear EST 15-MAY-2001
LOCUS 602649832F1 NIH_MGC_40 Homo sapiens CDNA clone IMAGE:4761054 5',
ACCESSION BG774910
VERSION BG774910
KEYWORDS BG774910.1 GI:14045227
SOURCE EST.
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NIH-MGC <http://mgc.nci.nih.gov/>.
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
<http://image.llnl.gov>
Plate: L1CM1612 row: o column: 07
High quality sequence stop: 874.
Location/Qualifiers

FEATURES

source 1..981
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4761054"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 60.3%; Score 36.2; DB 4; Length 981;
Best Local Similarity 77.2%; Pred. No. 0.4;
Matches 44; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGGTCGACTGCCCCCTCCGGCGCATGTTGACCTCAGCTCTGACACAGCCAGCC 57
 |||||
 Db 21 GGCTCCACCGCCCCCGCCAGCCCATGTTGTCACCTCGCCCGGACACAGCCCGCC 77
 |||||

RESULT 15

BQ923149 1536 bp mRNA linear EST 20-AUG-2002
 LOCUS BQ923149
 DEFINITION AGENCOURT 8929207 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6484568

5', mRNA sequence.

ACCESSION BQ923149
 VERSION BQ923149.1 GI:22338180
 KEYWORDS EST.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 1536)

NIH-MGC http://mgs.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov
 Tissue Procurement: DCTD/DTP

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LICM2670 row: p column: 09

High quality sequence stop: 287.

Location/Qualifiers

1..1536

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6484568"

/rissue_type="carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;

Site_2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Library constructed by

Ling Hong in the laboratory of Gerald M. Rubin (University

of California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH_MGC Library."

ORIGIN

Query Match 60.3%; Score 36.2; DB 5; Length 1536;

Best Local Similarity 77.2%; Pred. No. 0.4;

Matches 44; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGGTCGACTGCCCCCTCCGGCGCATGTTGACCTCAGCTCTGACACAGCCAGCC 57
 |||||
 Db 18 GGCTCCACCGCCCCCGCCAGCCCATGTTGTCACCTCGCCCGGACACAGCCCGCC 74
 |||||

Search completed: January 16, 2005, 02:55:38
 Job time : 1552.6 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 16:00:11 : Search time 170.5 Seconds
(without alignments)
1847.304 Million cell updates/sec

Title: US-10-057-136-10
Perfect score: 60
Sequence: 1 gggctgactgcgccctccggc.....ctgacacaagcgccagcccca 60

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : N_Geneseq_23Sep04:*
- 1: geneseqn1980s:*
 - 2: geneseqn1990s:*
 - 3: geneseqn2000s:*
 - 4: geneseqn2001as:*
 - 5: geneseqn2001bs:*
 - 6: geneseqn2002as:*
 - 7: geneseqn2002bs:*
 - 8: geneseqn2003as:*
 - 9: geneseqn2003bs:*
 - 10: geneseqn2003cs:*
 - 11: geneseqn2003ds:*
 - 12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	2 AAV48323	Aav48323 Nucleotid
2	39.4	65.7	4144	12 ADI57666	Adi57666 Human bre
3	38.2	63.7	60	2 AAV48316	Aav48316 Nucleotid
4	38.2	63.7	309	1 AAN90579	Aan90579 PDF9.3 CD
5	38.2	63.7	1194	12 ADI57712	Adi57712 Human bre
6	38.2	63.7	1378	12 ADI57693	Adi57693 Human bre
7	38.2	63.7	1424	12 ADO23180	Ado23180 Antisense
8	38.2	63.7	1428	6 ABL60159	Abi60159 Human MUC
9	38.2	63.7	1457	12 ADO23125	Ado23125 Human MUC
10	38.2	63.7	1457	12 ADF32627	Adf32627 Plasmid J
11	38.2	63.7	1527	2 AAV48329	Aav48329 MiniMUC1
12	38.2	63.7	1572	5 AAS00585	Aas00585 Human MUC
13	38.2	63.7	1614	12 ADK70370	Adk70370 Respirato
14	38.2	63.7	1630	12 ADI57708	Adi57708 Human bre
15	38.2	63.7	1634	12 ADI57689	Adi57689 Human bre
16	38.2	63.7	1712	12 ADI57686	Adi57686 Human bre
17	38.2	63.7	1738	12 ADI57669	Adi57669 Human bre
18	38.2	63.7	1755	12 ADI57673	Adi57673 Human bre
19	38.2	63.7	1774	12 ADE43991	Ade43991 Plasmid J
20	38.2	63.7	1774	12 ADF32625	Adf32625 Plasmid J
21	38.2	63.7	1799	12 ADO23124	Ado23124 Human MUC

22	38.2	63.7	1803	12 ADI57699	Adi57699 Human bre
23	38.2	63.7	1804	6 ABL67539	Abi67539 Thyroid c
24	38.2	63.7	1804	9 AAD56950	Aad56950 Human muc
25	38.2	63.7	1804	10 ADD14719	Add14719 Human src
26	38.2	63.7	1804	12 ADP13294	Adp13294 Renal cel
27	38.2	63.7	1805	12 ADO28642	Ado28642 Human MUC
28	38.2	63.7	1808	12 ADI57706	Adi57706 Human bre
29	38.2	63.7	1818	12 ADF32633	Adf32633 Plasmid J
30	38.2	63.7	1823	6 ABZ35228	Abz35228 Human gen
31	38.2	63.7	1823	12 ADI57707	Adi57707 Human bre
32	38.2	63.7	1835	12 ADF32631	Adf32631 Plasmid J
33	38.2	63.7	1874	12 ADI57688	Adi57688 Human bre
34	38.2	63.7	1882	12 ADI57677	Adi57677 Human bre
35	38.2	63.7	1918	12 ADI57672	Adi57672 Human bre
36	38.2	63.7	1930	12 ADI57678	Adi57678 Human bre
37	38.2	63.7	1945	12 ADI57676	Adi57676 Human bre
38	38.2	63.7	1949	12 ADI57698	Adi57698 Human bre
39	38.2	63.7	1953	12 ADI57668	Adi57668 Human bre
40	38.2	63.7	2045	12 ADI57701	Adi57701 Human bre
41	38.2	63.7	2049	12 ADI57682	Adi57682 Human bre
42	38.2	63.7	2090	12 ADI57705	Adi57705 Human bre
43	38.2	63.7	2094	12 ADI57681	Adi57681 Human bre
44	38.2	63.7	2135	12 ADF32629	Adf32629 Plasmid J
45	38.2	63.7	2194	12 ADI57683	Adi57683 Human bre

ALIGNMENTS

RESULT 1
ID AAV48323 standard; DNA; 60 BP.
XX AAV48323;
AC
XX 20-NOV-1998 (first entry)
DT
XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit R8.
XX
XX ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW tumour; tumour-associated antigen.
XX
OS Homo sapiens.
XX
XX PN W09837095-A2.
XX
XX PD 27-AUG-1998.
XX
XX PF 24-FEB-1998; 98WO-US003693.
XX
XX PR 24-FEB-1997; 97US-0038253P.
XX
XX (THER-) THERION BIOLOGICS CORP.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (DAND) DANA FARBER CANCER INST INC.
XX
XX Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
DR WPI, 1998-467492/40.
XX
XX New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
XX Disclosure; Page 11; 42pp; English.
XX
XX The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens

```
XX
SQ Sequence 60 BP; 10 A; 24 C; 17 G; 9 T; 0 U; 0 Other;
Query Match
Best Local Similarity 100.0%; Score 60; DB 2; Length 60;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGTCGACTGCCCTCCGGCGCATGTGTGACCTCAGCTCTGTACACAGGCCAGCCCCA 60
Db 1 GGGTCGACTGCCCTCCGGCGCATGTGTGACCTCAGCTCTGTACACAGGCCAGCCCCA 60
RESULT 2
ADIS7666
ID ADIS7666 standard; cDNA; 4144 BP.
XX
AC ADIS7666;
XX
DT 22-APR-2004 (first entry)
XX
DE Human breast specific nucleic acid (BSNA) #37.
XX
KW Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
KW breast cancer; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2003106648-A2.
XX
PD 24-DEC-2003.
XX
PF 16-JUN-2003; 2003WO-US018934.
XX
PR 14-JUN-2002; 2002US-0389327P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
XX
DR WPI; 2004-082185/08.
DR P-PSDB; ADIS7740.
XX
PT Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.
XX
PS Claim 1; SEQ ID NO 37; 370pp; English.
XX
CC The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the
CC hybridisation. The nucleic acids are useful for determining the
CC of a BSP in a sample which involves contacting the sample with suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein in
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.
XX
SQ Sequence 4144 BP; 634 A; 1896 C; 1058 G; 552 T; 0 U; 4 Other;
Query Match
65.7%; Score 39.4; DB 12; Length 4144;
```

```
Best Local Similarity 78.0%; Pred. No. 0.0022;
Matches 46; Conservative 1; Mismatches 12; Indels 0; Gaps 0;
QY 1 GGGTCGACTGCCCTCCGGCGCATGTGTGACCTCAGCTCTGTACACAGGCCAGCCCC 59
Db 693 GGCTCCACCGCCCCCGCCAGCCAGGTGTACCTCGGCCCGGACACAGCGCGCCCC 751
RESULT 3
AAV48316
ID AAV48316 standard; cDNA; 60 BP.
XX
AC AAV48316;
XX
DT 20-NOV-1998 (first entry)
XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit.
XX
KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW tumour; tumour-associated antigen.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH CDS 1..60
FT /*tag= a
FT /product= "MUC1 tandem repeat unit"
XX
PN WO9837095-A2.
XX
PD 27-AUG-1998.
XX
PF 24-FEB-1998; 98WO-US003693.
XX
PR 24-FEB-1997; 97US-0038253P.
XX
PA (THER-) THERION BIOLOGICS CORP.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA (DAND ) DANA FARBER CANCER INST INC.
XX
PI Schlom J, Kantar J, Kufe D, Panicali D, Gritz L;
XX
DR WPI; 1998-467492/40.
DR P-PSDB; AAW77229.
XX
PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS Example 1; Page 20; 42pp; English.
XX
CC The MUC1 tandem repeat unit was used to create an immunogenic mini-MUC1
CC fragment for inclusion in a recombinant pox virus (RPV). The RPV was used
CC in a pharmaceutical composition also containing an immunomodulator to
CC generate MUC1 specific cytotoxic T-lymphocytes. The recombinant pox virus
CC therefore encodes an immunogenic MUC1 fragment that does not undergo
CC significant genetic deletion, thereby providing an unexpectedly stable
CC and immunogenic pox virus. They can be used to prevent or treat tumours
CC expressing MUC1 tumour-associated antigens
XX
SQ Sequence 60 BP; 7 A; 33 C; 16 G; 4 T; 0 U; 0 Other;
Query Match
63.7%; Score 38.2; DB 2; Length 60;
Best Local Similarity 78.0%; Pred. No. 0.0032;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 1 GGGTCGACTGCCCTCCGGCGCATGTGTGACCTCAGCTCTGTACACAGGCCAGCCCC 59
Db 1 GGCTCCACCGCCCCCGCCAGCCAGGTGTACCTCGGCCCGGACACAGCGCGCCCC 59
RESULT 4
AAAN90579/C
ID AAAN90579 standard; cDNA; 309 BP.
```

```

XX AC AAN90579;
XX DT 27-AUG-2003 (revised)
XX DT 25-MAR-2003 (revised)
XX DT 04-DEC-1989 (first entry)
XX DE PDF9.3 cDNA insert.
XX KM PDF9.3; human DF3 breast carcinoma-associated antigen epitope.
XX OS Human MCF-7 breast carcinoma cells; 'ATCC HTB22'.
XX PN WO8907107-A.
XX PD 10-AUG-1989.
XX PF 29-JAN-1988; 88US-00149831.
XX PR 29-JAN-1988; 88US-00149831.
XX PA (DANA-) DANA-FARBER CANCER.
XX PI Kufe DW;
XX DR WPI; 1989-248989/34.
XX DR P-PSDB; AAP91045, AAP91053, AAP91054, AAP90146.
XX PT Recombinant polypeptide(s) - contains DF 3 breast carcinoma antigen
XX PT epitope and useful as assay reagents, and encoding DNA sequences.
XX PS Claim 1; Fig 4; 31pp; English.
XX CC The sequence encodes a carbohydrate-free polypeptide contg. DF3 breast
XX CC carcinoma antigen epitope. Useful as a competitive binding assay reagent
XX CC and improves diagnosis. The cDNA consists of nearly identical 60 BP
XX CC tandem repeats which are 85% GC-rich. See also AAP91045, AAP91046,
XX CC AAP91053, AAP91054. (Updated on 25-MAR-2003 to correct PA field.)
XX CC (Updated on 27-AUG-2003 to correct OS field.)
XX SQ Sequence 309 BP; 21 A; 92 C; 160 G; 36 T; 0 U; 0 Other;

Query Match          63.7%; Score 38.2; DB 1; Length 309;
Best Local Similarity 78.0%; Pred. No. 0.0039;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGGTCGACTGCCCTCCGGCGCATGGTGTGACCTCAGCTCCTGACACAAGGCCAGCCCC 59
Db 266 GGCTCCACCGCGCCCCCGACGCCGAGTGTACCTCGGCCCGGACACCAAGCGCGCCCC 208

RESULT 5
ADIS7712
ID ADIS7712 standard; cDNA; 1194 BP.
XX AC ADIS7712;
XX DT 22-APR-2004 (first entry)
XX DE Human breast specific nucleic acid (BSNA) #83.
XX KM Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
XX KM breast cancer; cytostatic.
XX OS Homo sapiens.
XX PN WO2003106648-A2.
XX PD 24-DEC-2003.
XX PF 16-JUN-2003; 2003WO-US018934.
XX PR 14-JUN-2002; 2002US-0389327P.

```

```

XX PA (DIAD-) DIADEXUS INC.
XX PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
XX DR WPI; 2004-082185/08.
XX DR P-PSDB; ADIS7782.
XX PT Novel isolated polypeptide comprising breast specific protein sequences,
XX PT useful for diagnosing or monitoring presence and metastases of breast
XX PT cancer in patient.
XX PS Claim 1; SEQ ID NO 83; 370pp; English.
XX CC The invention relates to human breast specific nucleic acids (BSNA) and
XX CC the breast specific proteins (BSP) they encode. The nucleic acids are
XX CC useful for determining the presence of a BSNA in a sample which involves
XX CC contacting the sample with a BSNA under conditions in which the BSNA will
XX CC selectively hybridise to a BSNA in the sample, and detecting the
XX CC hybridisation. The nucleic acids are useful for determining the presence
XX CC of a BSP in a sample which involves contacting the sample with suitable
XX CC reagent under conditions in which the reagent will selectively interact
XX CC with the BSP, and detecting the interaction of the reagent with a BSP in
XX CC the sample. The nucleic acids and proteins are useful for diagnosing or
XX CC monitoring the presence and metastases of breast cancer in a patient,
XX CC which involves determining an amount of nucleic acid or protein and
XX CC comparing the determined amount of nucleic acid or protein in the sample
XX CC of the patient to the amount of a breast specific marker in a normal
XX CC control, where a difference in the determined amount in the sample
XX CC compared to the amount in the control is associated with the presence of
XX CC breast cancer. The sequences are useful for treating a patient with
XX CC breast cancer, involving administering a composition consisting of a BSNA
XX CC or a BSP to a patient, where the administration induces an immune
XX CC response against the breast cancer cell expressing the BSNA or BSP. This
XX CC sequence represents a human BSNA of the invention.
XX SQ Sequence 1194 BP; 243 A; 438 C; 275 G; 238 T; 0 U; 0 Other;

Query Match          63.7%; Score 38.2; DB 12; Length 1194;
Best Local Similarity 78.0%; Pred. No. 0.0047;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGGTCGACTGCCCTCCGGCGCATGGTGTGACCTCAGCTCCTGACACAAGGCCAGCCCC 59
Db 596 GGCTCCACCGCGCCCCCGACGCCGAGTGTACCTCGGCCCGGACACCAAGCGCGCCCC 654

RESULT 6
ADIS7693
ID ADIS7693 standard; cDNA; 1378 BP.
XX AC ADIS7693;
XX DT 22-APR-2004 (first entry)
XX DE Human breast specific nucleic acid (BSNA) #64.
XX KM Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
XX KM breast cancer; cytostatic.
XX OS Homo sapiens.
XX PN WO2003106648-A2.
XX PD 24-DEC-2003.
XX PF 16-JUN-2003; 2003WO-US018934.
XX PR 14-JUN-2002; 2002US-0389327P.
XX PA (DIAD-) DIADEXUS INC.
XX PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;

```


XX WPI; 2004-082185/08.
DR P-PSDB; ADI57765.

PT Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.

PS Claim 1; SEQ ID NO 64; 370pp; English.

XX The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the
CC hybridisation. The nucleic acids are useful for determining the presence
CC of a BSP in a sample which involves contacting the sample with suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein and
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.

XX Sequence 1378 BP; 295 A; 489 C; 316 G; 278 T; 0 U; 0 Other;

Query Match 63.7%; Score 38.2; DB 12; Length 1378;
Best Local Similarity 78.0%; Pred. No. 0.0047;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGGTCGACTGCCCTCCGGCGCATGTGTGACCTCAGCTCTGACACAGGCCAGCCCC 59
Db 596 GGCTCCACCGCCCCCGCCAGCCCAAGGTGTCACTCGGCCCGGACACAGCGCGGCC 654

RESULT 7

ID ADO23180 standard; RNA; 1424 BP.

AC ADO23180;

DT 12-AUG-2004 (first entry)

DE Antisense human MUC1 mucin glycoprotein RNA (coding sequence) SeqID 75.

KW human; MUC1; mucin glycoprotein; cancer; chemotherapeutic; MUC1/ECD; ss;

OS Homo sapiens.

EN WO2004044160-A2.

PD 27-MAY-2004.

PF 12-NOV-2003; 2003WO-US035848.

PR 13-NOV-2002; 2002US-00293391.

PR 29-MAY-2003; 2003US-00447839.

PA (DAND) DANA FARBER CANCER INST INC.

PA (ILEX-) ILEX PROD INC.

PI Kuife DW, Kharbanda S, Weitman SD;
XX WPI; 2004-420304/39.

XX Double-stranded RNA complex useful for inhibiting proliferation of cancer
PT cell expressing MUC1 mucin glycoprotein, comprises first and second RNA
PT sequences.

PS Disclosure; SEQ ID NO 75; 112pp; English.

XX This invention relates to novel modulators of the human MUC1 mucin
CC glycoprotein for use in cancer therapeutics, where MUC1 is a protein that
CC acts to inhibit the apoptotic response to genotoxic stress caused by
CC chemotherapeutic agents. In particular, it refers to modulators of the
CC MUC1 extracellular domain (MUC1/ECD). The method refers to using double-
CC stranded RNA complexes as MUC1 interference RNA compositions such that
CC MUC1 expression is inhibited, which in turn inhibits cancer cell
CC proliferation. The present invention describes screening assays to
CC identify compounds that inhibit the binding of various MUC1 ligands such
CC as neuroligin 2, as well as agonists, antagonists and antibodies thereof.
CC Furthermore, it provides MUC1 antisense, small interfering RNA (siRNA)
CC and small molecules in combination with chemotherapeutic agents that are
CC useful in the field of cancer therapy. This polynucleotide sequence is
CC the antisense human MUC1 RNA of the invention.

XX Sequence 1424 BP; 309 A; 325 C; 495 G; 0 T; 295 U; 0 Other;

Query Match 63.7%; Score 38.2; DB 12; Length 1424;
Best Local Similarity 78.0%; Pred. No. 0.0048;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGGTCGACTGCCCTCCGGCGCATGTGTGACCTCAGCTCTGACACAGGCCAGCCCC 59
Db 1041 GGCTCCACCGCCCCCGCCAGCCCAAGGTGTCACTCGGCCCGGACACAGCGCGGCC 983

RESULT 8

ABL60159
ID ABL60159 standard; cDNA; 1428 BP.

AC ABL60159;

DT 22-JUL-2002 (first entry)

DE Human MUC1 encoding cDNA SEQ ID NO 2.

KW Human; mucin 1; MUC1; transmembrane protein; SNP; cancer; cytostatic;

KW single nucleotide polymorphism; haplotyping; genotyping; drug;

KW antiinflammatory; gene; ss.

OS Homo sapiens.

FT Key

FT CDS

FT variation

FT

FT

FT

FT

FT

FT

FT

FT

Location/Qualifiers

1..1428

/*tag= a

/product= "MUC1"

replace(1009,A)

/*tag= b

/standard_name= "single nucleotide polymorphism"

/note= "SNP allelic variation results in Val substituted

by Met at position 337 of the MUC1 protein (ABB77476)"

WO200226765-A2.

04-APR-2002.

25-SEP-2001; 2001WO-US030151.

28-SEP-2000; 2000US-0236113P.

(GENA-) GENAISSANCE PHARM INC.

Chew A, Koshiy B;

WPI; 2002-405042/43.

P-PSDB; ABB77476.

```

XX New genetic variants of mucin 1, Transmembrane gene, useful in studying
PT expression and function of protein encoded by the gene and for screening
PT drugs to treat diseases e.g. cancer.
XX
XX Claim 23; Fig 2; 75pp; English.
XX
CC The invention relates to a polynucleotide (ABU60158, ABU60159) encoding
CC mucin 1/MUC1 (ABB77476), Transmembrane isogene. The invention describes
CC novel genetic variants of the MUC1 gene. The invention is useful for
CC haplotyping/genotyping the MUC1 gene in an individual and identifying an
CC association between a trait and at least one of the haplotypes or
CC haplotype pairs of MUC1 gene. MUC1 is useful for studying the expression
CC and function of MUC1 and expressing MUC1 protein for use in screening for
CC candidate drugs to treat diseases related to MUC1 activity and in
CC studying the effect of the variation on the biological activity of MUC1
CC as well as on the binding affinity of candidate drugs targeting MUC1 for
CC the treatment of e.g. cancer. MUC1 is further used by the pharmaceutical
CC research scientist to validate MUC1 as a candidate target for and in
CC design of clinical trials of candidate drugs for, treating a specific
CC condition drugs or disease predicted to be associated with MUC1 activity.
CC MUC1 antibodies are useful in a variety of diagnostic and prognostic
CC formats and therapeutic methods
XX
SQ Sequence 1428 BP; 296 A; 496 C; 327 G; 309 T; 0 U; 0 Other;

Query Match          63.7%; Score 38.2; DB 6; Length 1428;
Best Local Similarity 78.0%; Pred. No. 0.0048;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

OY      1 GGGTCGACTGCCCTCCGGCGCATGCTGTGACCTCCTCTGACACAAGGCCAGCCCC 59
         |||||
DB      385 GGCTCCACCGCCCCCGCCAGCCAGCGTGTCACTCGGCCCGGACACCAAGCGCGCCCC 443

RESULT 9
ADO23125
ID      ADO23125 standard; RNA; 1428 BP.
XX
XX      ADO23125;
AC
XX
DT      12-AUG-2004 (first entry)
XX
DE      Human MUC1 mucin glycoprotein RNA (coding sequence) SeqID 20.
XX
XX      human; MUC1; mucin glycoprotein; cancer; chemotherapeutic; MUC1/ECD; ss.
XX
XX      Homo sapiens.
OS
XX
XX      WO2004044160-A2.
PN
XX
XX      27-MAY-2004.
PD
XX
XX      12-NOV-2003; 2003WO-US035848.
PF
XX
XX      13-NOV-2002; 2002US-00293391.
PR      29-MAY-2003; 2003US-00447839.
XX
XX      (DAND ) DANA FARBER CANCER INST INC.
PA      (ILEX-) ILEX PROD INC.
XX
XX      Kuife DW, Kharbada S, Weltman SD;
PI
XX      WPI; 2004-420304/39.
DR
XX
XX      Double-stranded RNA complex useful for inhibiting proliferation of cancer
PT      cell expressing MUC1 mucin glycoprotein, comprises first and second RNA
PT      sequences.
XX
XX      Claim 2; SEQ ID NO 20; 112pp; English.
PS
XX
XX      This invention relates to novel modulators of the human MUC1 mucin
CC      glycoprotein for use in cancer therapeutics, where MUC1 is a protein that

```

```

CC acts to inhibit the apoptotic response to genotoxic stress caused by
CC chemotherapeutic agents. In particular, it refers to modulators of the
CC MUC1 extracellular domain (MUC1/ECD). The method refers to using double-
CC stranded RNA complexes as MUC1 interference RNA compositions such that
CC MUC1 expression is inhibited, which in turn inhibits cancer cell
CC proliferation. The present invention describes screening assays to
CC identify compounds that inhibit the binding of various MUC1 ligands such
CC as neuregulin 2, as well as agonists, antagonists and antibodies thereof.
CC Furthermore, it provides MUC1 antisense, small interfering RNA (siRNA)
CC and small molecules in combination with chemotherapeutic agents that are
CC useful in the field of cancer therapy. This polynucleotide sequence is
CC the human MUC1 RNA of the invention.
XX
SQ Sequence 1428 BP; 296 A; 496 C; 327 G; 0 T; 309 U; 0 Other;

Query Match          63.7%; Score 38.2; DB 12; Length 1428;
Best Local Similarity 71.2%; Pred. No. 0.0048;
Matches 42; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

OY      1 GGGTCGACTGCCCTCCGGCGCATGCTGTGACCTCCTCTGACACAAGGCCAGCCCC 59
         |||||
DB      385 GGCTCCACCGCCCCCGCCAGCCAGCGUGUACCTCGGCCCGGACACCAAGCGCGCCCC 443

RESULT 10
ADF32627
ID      ADF32627 standard; DNA; 1457 BP.
XX
XX      ADF32627;
AC
XX
XX      26-FEB-2004 (first entry)
DT
XX
DE      Plasmid JNW358 MUC-1 nucleotide sequence.
XX
XX      MUC-1 antigen; immune response; MUC-1; variable number of tandem repeat;
XX      VNTR; repeat unit; tumour; metastasis; cytostatic; vaccine; gene therapy;
XX      gene; ds.
XX
XX      Synthetic.
OS
XX
XX      WO2003100060-A2.
PN
XX
XX      04-DEC-2003.
PD
XX
XX      23-MAY-2003; 2003WO-EP005594.
PF
XX
XX      24-MAY-2002; 2002GB-00012046.
PR
XX
XX      (GLAX ) GLAXO GROUP LTD.
PA
XX
XX      Burden N, Ellis JH, Hamblin PA;
PI
XX
XX      WPI; 2004-042811/04.
DR
XX
XX      New nucleic acid molecule encoding a MUC-1 antigen, useful for preparing
PT      a composition for treating or preventing tumors or metastases.
XX
XX      Example; Fig 3; 66pp; English.
XX
XX      The present invention describes a nucleic acid molecule which encodes a
CC      MUC-1 antigen. The nucleic acid is capable of raising an immune response
CC      in vivo, has reduced susceptibility to recombination than full-length MUC
CC      -1 and comprises between 1 and 15 variable number of tandem repeats
CC      (VNTR) perfect repeat units. Also described: (1) a plasmid comprising the
CC      DNA molecule; (2) a protein encoded by the nucleic acid; (3) a
CC      pharmaceutical composition comprising the nucleic acid; plasmid or
CC      protein and an excipient, diluent or carrier; and (4) a method of
CC      treating or preventing tumours or metastases. A MUC1 antigen has
CC      cytostatic activity, and can be used in vaccines, and in gene therapy.
CC      The nucleic acid is useful for preparing a composition for treating or
CC      preventing tumours or metastases. The present sequence is used in the
CC      exemplification of the present invention.
XX

```

Sequence 1457 BP; 302 A; 506 C; 334 G; 315 T; 0 U; 0 Other;

Query Match
Best Local Similarity 63.7%; Score 38.2; DB 12; Length 1457;
Matches 46; Conservativity 78.0%; Pred. No. 0.0048;

Matches 46; Conservativity 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGGTCGACTGCCCCCTCCGCGCATGTGTGACCTCAGCTCTCTGACACAAAGGCCAGCCCC 59
DB 403 GGCTCCACCGCCCCCGCCAGCCACGGTGTCACTCGGCCCCCGGACACCAAGCGCGCCCC 461

RESULT 11

AAV48329
ID AAV48329 standard; cDNA; 1527 BP.

XX AC AAV48329;

DT 20-NOV-1998 (first entry)

XX DE MiniMUC1 gene.

XX KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
XX KW tumour; tumour-associated antigen.

OS Homo sapiens.

Key Location/Qualifiers
CDS 1..1527
/*tag= a
/product= "MiniMUC1 protein"

XX PN WO9837095-A2.
XX PD 27-AUG-1998.
XX PF 24-FEB-1998; 98WO-US003693.
XX PR 24-FEB-1997; 97US-0038253P.

XX PA (THER-) THERION BIOLOGICS CORP.
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX PA (DAND) DANA FARBER CANCER INST INC.

PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX WPI; 1998-467492/40.
DR P-PSDB; AAW77233.

XX PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
XX PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX PS Example 1; Page 21-22; 42pp; English.

XX CC The immunogenic mini-MUC1 fragment was created from MUC1 tandem repeat
XX CC units for inclusion in a recombinant pox virus (RPV). The RPV was used in
XX CC a pharmaceutical composition also containing an immunomodulator to
XX CC generate MUC1 specific cytotoxic T-lymphocytes. The recombinant pox virus
XX CC therefore encodes an immunogenic MUC1 fragment that does not undergo
XX CC significant genetic deletion, thereby providing an unexpectedly stable
XX CC and immunogenic pox virus. They can be used to prevent or treat tumours
XX CC expressing MUC1 tumour-associated antigens

XX SQ Sequence 1527 BP; 296 A; 573 C; 351 G; 307 T; 0 U; 0 Other;

Query Match
Best Local Similarity 63.7%; Score 38.2; DB 2; Length 1527;
Matches 46; Conservativity 78.0%; Pred. No. 0.0048;

QY 1 GGGTCGACTGCCCCCTCCGCGCATGTGTGACCTCAGCTCTCTGACACAAAGGCCAGCCCC 59
DB 226 GGCTCCACCGCCCCCGCCAGCCACGGTGTCACTCGGCCCCCGGACACCAAGCGCGCCCC 284

RESULT 12

AA500585
ID AA500585 standard; DNA; 1572 BP.

XX AC AA500585;

DT 12-SEP-2001 (first entry)

XX DE Human MUC1 DNA sequence.

XX KW Human; MUC1; antigenic peptide; major histocompatibility complex; MHC-I;
XX KW glycoprotein; cytotoxic T lymphocytes; T cell response; cancer; vaccine;
XX KW cancer gene therapy; diagnosis; treatment; inflammatory disorder; ds;
XX KW organ transplant rejection; graft versus host disease.

OS Homo sapiens.

Key Location/Qualifiers
CDS 58..1545
/*tag= a
/product= "Human MUC1"

XX PN WO200118035-A2.

XX PD 15-MAR-2001.

XX PF 07-SEP-2000; 2000WO-EP008761.

XX PR 08-SEP-1999; 99GB-00021242.
XX PR 10-SEP-1999; 99EP-00402237.
XX PR 03-MAR-2000; 2000US-0187215P.

XX PA (TRGE) TRANSGENE SA.
XX PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.

PI Taylor-Papadimitriou J, Heukamp LC, Offringa R, Melief CJM;
XX PI Acres B, Thomas M;

XX DR WPI; 2001-235187/24.
XX DR P-PSDB; AAU00539.

XX PT New antigenic polypeptides of MUC-1 protein which activate cytotoxic T
XX PT lymphocyte proteins and their analogs, useful for identifying a major
XX PT histocompatibility complex class I restricted T cell response and for
XX PT diagnosing cancer.

XX PS Claim 1; Page 66-68; 81pp; English.

XX CC The sequence represents a polynucleotide which encodes MUC1 polypeptide.
XX CC Derivative antigenic peptides of this protein bind at least one major
XX CC histocompatibility complex class I (MHC-I) glycoprotein, which activates
XX CC cytotoxic T lymphocytes to induce a protective response against tumours.
XX CC Diagnosis of cancer involves determining the presence or absence in a
XX CC host cell of a MHC class I restricted T cell response to a MUC1
XX CC derivative, where the presence of the MHC class I restricted T cell
XX CC response indicates that the host has cancer. Measurement of the level of
XX CC MHC class I restricted T cell response is also useful to monitor the
XX CC severity of cancer, a larger response indicating a more severe cancer.
XX CC MUC1 derivatives are useful in cancer therapy and to follow MUC1 specific
XX CC immune responses in patients during the course of disease and/or
XX CC treatment. MUC1 DNA is useful in cancer gene therapy, vaccination and
XX CC diagnosis. Compositions of the sequences are used in vaccines and
XX CC as an inflammatory disorder, organ transplant rejection or graft versus
XX CC host disease

XX SQ Sequence 1572 BP; 312 A; 567 C; 364 G; 329 T; 0 U; 0 Other;

Query Match
Best Local Similarity 63.7%; Score 38.2; DB 5; Length 1572;
Matches 46; Conservativity 78.0%; Pred. No. 0.0048;

QY 1 GGGTCGACTGCCCCCTCCGCGCATGTGTGACCTCAGCTCTCTGACACAAAGGCCAGCCCC 59

Db 442 GGCTCCACCGCCCCCGCCGACGGTGTCACTCGCCCCCGACACCAAGCGCCCCC 500

RESULT 13

ADK70370

ID ADK70370 standard; cDNA; 1614 BP.

XX ADK70370;

DT 06-MAY-2004 (first entry)

DE Respiratory disease differentially expressed cDNA #106.

KW ds; gene; cytostatic; respiratory; antiasthmatic; gene therapy;

KM differential gene expression; respiratory disorder; lung cancer;

KW chronic obstructive pulmonary disease; emphysema; asthma.

XX Homo sapiens.

OS WO2003101283-A2.

PN 11-DEC-2003.

PD 02-JUN-2003; 2003WO-US017409.

PF 04-JUN-2002; 2002US-0386005P.

XX (INCY-) INCYTE CORP.

PA Rickett PK, Krasnow R;

XX WPI; 2004-042945/04.

DR New combination comprising cDNAs and proteins that are differentially

XX expressed in respiratory disorders, useful for diagnosing or treating

PT respiratory diseases e.g. lung cancer, chronic obstructive pulmonary

PT diseases or asthma.

XX Claim 1; SEQ ID NO 106; 343pp; English.

CC The invention relates to cDNA sequences that are differentially expressed

CC in respiratory disorders or their complements or encoded proteins. The

CC cDNAs and proteins are useful for diagnosing, treating or monitoring

CC treatment of a subject with a respiratory disease including lung cancer,

CC chronic obstructive pulmonary diseases, emphysema or asthma. The protein

CC is also useful for screening molecules or compounds to identify at least

CC one ligand which specifically binds the protein. It is also useful for

CC preparing and purifying a polyclonal or monoclonal antibody. This

CC sequence corresponds to a cDNA of the invention.

XX Sequence 1614 BP; 329 A; 556 C; 374 G; 353 T; 0 U; 2 Other;

Query Match 63.7%; Score 38.2; DB 12; Length 1614;

Best Local Similarity 78.0%; Pred. No. 0.0048;

Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

OY 1 GGCTGACTGCCCTCCGGCGCATGTGTGACTCAGCTCCTGACACAAGGCCAGCCCC 59

DB 434 GGCTCCACCGCCCCCGCCGACGGTGTCACTCGCCCCCGACACCAAGCGCCCCC 492

RESULT 14

AD157708

ID AD157708 standard; cDNA; 1630 BP.

XX AD157708;

DT 22-APR-2004 (first entry)

DE Human breast specific nucleic acid (BSNA) #79.

XX Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;

KM breast cancer; cytostatic.

XX Homo sapiens.

OS WO2003106648-A2.

PN 24-DEC-2003.

PD 16-JUN-2003; 2003WO-US018934.

PF 14-JUN-2002; 2002US-0389327P.

XX (DIAD-) DIADEXUS INC.

PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;

XX WPI; 2004-082185/08.

DR P-PSDB; AD157779.

XX Novel isolated polypeptide comprising breast specific protein sequences,

PT useful for diagnosing or monitoring presence and metastases of breast

PT cancer in patient.

XX Claim 1; SEQ ID NO 79; 370pp; English.

CC The invention relates to human breast specific nucleic acids (BSNA) and

CC the breast specific proteins (BSP) they encode. The nucleic acids are

CC useful for determining the presence of a BSNA in a sample which involves

CC contacting the sample with a BSNA under conditions in which the BSNA will

CC selectively hybridise to a BSNA in the sample, and detecting the

CC hybridisation. The nucleic acids are useful for determining the presence

CC of a BSP in a sample which involves contacting the sample with suitable

CC reagent under conditions in which the reagent will selectively interact

CC with the BSP, and detecting the interaction of the reagent with a BSP in

CC the sample. The nucleic acids and proteins are useful for diagnosing or

CC monitoring the presence and metastases of breast cancer in a patient,

CC which involves determining an amount of nucleic acid or protein and

CC comparing the determined amount of nucleic acid or protein in the sample

CC of the patient to the amount of a breast specific marker in a normal

CC control, where a difference in the determined amount in the sample

CC compared to the amount in the control is associated with the presence of

CC breast cancer. The sequences are useful for treating a patient with

CC breast cancer, involving administering a composition consisting of a BSNA

CC or a BSP to a patient, where the administration induces an immune

CC response against the breast cancer cell expressing the BSNA or BSP. This

CC sequence represents a human BSNA of the invention.

XX Sequence 1630 BP; 322 A; 584 C; 405 G; 319 T; 0 U; 0 Other;

Query Match 63.7%; Score 38.2; DB 12; Length 1630;

Best Local Similarity 78.0%; Pred. No. 0.0048;

Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

OY 1 GGCTGACTGCCCTCCGGCGCATGTGTGACTCAGCTCCTGACACAAGGCCAGCCCC 59

DB 596 GGCTCCACCGCCCCCGCCAGCCAGGTGTCACTCGCCCCCGACACCAAGCGCCCCC 654

RESULT 15

AD157689

ID AD157689 standard; cDNA; 1634 BP.

XX AD157689;

DT 22-APR-2004 (first entry)

DE Human breast specific nucleic acid (BSNA) #60.

XX Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;

KM breast cancer; cytostatic.

XX Homo sapiens.

PN WO2003106648-A2.
XX
PD 24-DEC-2003.
XX
PF 16-JUN-2003; 2003WO-US018934.
XX
PR 14-JUN-2002; 2002US-0389327P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
XX WPI; 2004-082185/08.
DR P-PSDB; ADI57761.
XX
PT Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
XX cancer in patient.
PS Claim 1; SEQ ID NO 60; 370pp; English.
XX
CC The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the
CC hybridisation. The nucleic acids are useful for determining the presence
CC of a BSP in a sample which involves contacting the sample with the presence
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein and
CC comparing the determined amount of nucleic acid or protein in a normal
CC of the patient to the amount of a breast specific marker in the sample
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.
XX
SQ Sequence 1634 BP; 349 A; 541 C; 405 G; 337 T; 0 U; 2 Other;

Query Match 63.7%; Score 38.2; DB 12; Length 1634;
Best Local Similarity 78.0%; Pred. No. 0.0048;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCAGCTCCTGACACAGGCCACCC 59
Db 596 GGCTCCACCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC 654

Search completed: January 15, 2005, 20:36:11
Job time : 170.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 17:40:52; Search time 722.8 Seconds
(without alignments)
3925.542 Million cell updates/sec

Title: US-10-057-136-10

Perfect score: 60
Sequence: 1 gggctgactgcacctccggc.....ctgacacaagccagcccca 60

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 4526729 segs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	2297	6	BD272907 A recombi
2	38.2	63.7	120	6	AX192396 Sequence
3	38.2	63.7	120	6	BD000571 Human pol
4	38.2	63.7	1414	12	AF423031 Synthetic
5	38.2	63.7	1455	6	CQ715242 Sequence
6	38.2	63.7	1457	6	AX959914 Sequence
7	38.2	63.7	1572	6	AX093798 Sequence
8	38.2	63.7	1774	6	AX959684 Sequence
9	38.2	63.7	1774	6	AX959912 Sequence
10	38.2	63.7	1804	6	AR492318 Sequence
11	38.2	63.7	1804	6	AX335367 Sequence
12	38.2	63.7	1804	9	HUMMUCA8
13	38.2	63.7	1834	12	AF423030 Synthetic
14	38.2	63.7	1835	6	AX959918 Sequence
15	38.2	63.7	2135	6	AX959916 Sequence
16	38.2	63.7	4139	6	CQ834017 Sequence
17	38.2	63.7	4139	6	AX334899 Sequence
18	38.2	63.7	4139	6	AX335372 Sequence
19	38.2	63.7	4139	6	AX336712 Sequence

20	38.2	63.7	4139	6	AX409474 Sequence
21	38.2	63.7	4139	6	AX440481 Sequence
22	38.2	63.7	4139	9	HUMPANMU
23	38.2	63.7	7188	9	AY463543 Homo sapi
24	38.2	63.7	8181	6	AX406624 Sequence
25	38.2	63.7	8181	9	HUMPEM
26	38.2	63.7	8186	6	AR492320 Sequence
27	38.2	63.7	133525	9	AL713999 Human DNA
28	36.6	61.0	120	6	AX192397 Sequence
29	36.6	61.0	120	6	BD000572 Human pol
30	36.6	61.0	572	6	AR492319 Sequence
31	36.6	61.0	572	9	HUMDF3AA
32	36.6	61.0	1721	6	CQ771290 Sequence
33	36.6	61.0	1721	6	AR492306 Sequence
34	36.6	61.0	1721	6	AX335860 Sequence
35	36.6	61.0	1721	6	AX440427 Sequence
36	36.6	61.0	1721	6	AX587588 Sequence
37	36.6	61.0	1721	9	HSTBYMA
38	36.6	61.0	1971	6	AX963157 Sequence
39	36.6	61.0	2037	6	AX963159 Sequence
40	36.6	61.0	2238	9	HSETA
41	36.4	60.7	525	6	BD225141 Medicinal
42	36.4	60.7	891	6	BD225147 Medicinal
43	36.2	60.3	491	9	M22739 Human episi
44	36.2	60.3	518	6	AR492402 Sequence
45	36.2	60.3	518	9	HUMEPISAL

ALIGNMENTS

RESULT 1	BD272907	2297 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD272907	A recombinant vector expressing multiple constitutimulatory molecules and uses thereof.			
DEFINITION	BD272907	A recombinant vector expressing multiple constitutimulatory molecules and uses thereof.			
ACCESSION	BD272907.1	GI:33082675			
VERSION	JP 2002531133-A/1.				
KEYWORDS	JP 2002531133-A/1.				
SOURCE	synthetic construct				
ORGANISM	artificial sequences.				
REFERENCE	1 (bases 1 to 2297)				
AUTHORS	Schlom,J., Hodge,J. and Panicali,D.				
TITLE	A recombinant vector expressing multiple constitutimulatory molecules and uses thereof				
JOURNAL	Patent: JP 2002531133-A 1 24-SEP-2002;				
COMMENT	THE UNITED STATES OF AMERICA, THERION BIOLOGICS CORP				
	OS Artificial Sequence				
	PN JP 2002531133-A/1				
	PD 24-SEP-2002				
	PF 12-NOV-1999 JP 2000586927				
	PR 09-DEC-1998 US 60/111582				
	PI JEFFREY SCHLOM, JAMES HODGE, DENNIS PANICALI				
	PC C12N15/02, A61K35/12, A61K35/74, A61K35/76, A61K38/00, A61K39/00,				
	PC A61K39/12,				
	PC A61K39/125, A61K39/21, A61K39/235, A61K39/245, A61K39/275, A61K39/				
	PC 29, A61K48/00,				
	PC A61P1/04, A61P31/04, A61P31/10, A61P31/12, A61P35/00, A61P37/02, PC				
	A61P37/06,				
	PC C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12N7/00, C12Q1/02, G01N33/				
	PC 53, C12N15/00,				
	PC C12N5/00, A61K37/02				
	CC Description of Artificial Sequence: VECTOR SEQUENCE FH	Key			
	Location/Qualifiers				
FT	source	1..2297			
FEATURES	Location/Qualifiers				
source	1..2297				
	/organism="Artificial Sequence".				
	/organism="synthetic construct"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:32630"				
ORIGIN					

Query Match 100.0%; Score 60; DB 6; Length 2297;
Best Local Similarity 100.0%; Pred. No. 2.8e-07;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCCTCGACACAGAGGCCAGCCCCA 60
|||||
Db 766 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCCTCGACACAGAGGCCAGCCCCA 825

RESULT 2
AX192396/c 120 bp DNA linear PAT 15-AUG-2001
LOCUS Sequence 2 from Patent EP1103623.
DEFINITION AX192396
ACCESSION AX192396
VERSION AX192396.1 GI:15210363
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
Taylor-Papadimitriou,J., Burchell,J. and Gendler,S.
Human mucin core protein: nucleic acid probes, peptide fragments
and antibodies thereto, and uses thereof in diagnostic and
therapeutic methods
Patent: EP 1103623-A 2 30-MAY-2001;
JOURNAL IMPERIAL CANCER RESEARCH TECHNOLOGY LIMITED (GB)
FEATURES
source 1. 120
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 63.7%; Score 38.2; DB 6; Length 120;
Best Local Similarity 78.0%; Pred. No. 1.1;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCCTCGACACAGGCCAGCCCC 59
|||||
Db 87 GGCTCCACCGCCCCCGAGCCGACGCTGCTCGCCCGGACACAGGCCGCCCC 29

RESULT 3
BD000571/c 120 bp DNA linear PAT 31-JAN-2002
LOCUS Human polymorphic epithelial mucin core protein and nucleic acid
DEFINITION encoding the protein.
ACCESSION BD000571
VERSION BD000571.1 GI:18623684
KEYWORDS JP 2000333675-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
Papadimitrov,J.T., Jendora,S. and Bachneru,J.
Human polymorphic epithelial mucin core protein and nucleic acid
encoding the protein
Patent: JP 2000333675-A 2 05-DEC-2000;
JOURNAL IMPERIAL CANCER RESEARCH TECHNOLOGY LTD
COMMENT OS Homo sapiens (human)
PN JP 2000333675-A/2
PD 05-DEC-2000
PR 26-APR-2000 JP 2000125724
PR 07-JAN-1987 GB 8700269,07-JAN-1987 GB 8700279 PR
22-APR-1987 US 041306,09-NOV-1987 GB 8726172 PI JOYCE
TAYLOR PAPADIMITROV,SANDRA JENDORA,JOY BACHNERU PC
C12N15/02,A61K38/00,A61K39/395,A61K39/395,A61K49/00,A61P35/00, PC
C07K14/47,
PC C07K16/44,C12N5/10,C12P21/08//C12N15/02,C12R1:91),(C12N5/10,
PC C12R1:91),

PC C12N15/00,A61K37/02,C12N5/00,(C12N15/00,C12R1:91),(C12N5/00,
PC C12R1:91)
CC
FH Key Location/Qualifiers
FT source 1. 120
/organism="Homo sapiens (human)"
location/Qualifiers
1. 120
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 63.7%; Score 38.2; DB 6; Length 120;
Best Local Similarity 78.0%; Pred. No. 1.1;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCCTCGACACAGGCCAGCCCC 59
|||||
Db 87 GGCTCCACCGCCCCCGAGCCGACGCTGCTCGCCCGGACACAGGCCGCCCC 29

RESULT 4
AF423031 1414 bp mRNA linear SYN 10-JUL-2003
LOCUS AF423031
DEFINITION Synthetic construct Homo sapiens mucin variant MUC1-CT58 (MUC1)
ACCESSION AF423031
VERSION AF423031.1 GI:19338621
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 1414)
Hinojosa-Kurtzberg,A.M., Johansson,M.E., Madsen,C.S., Hansson,G.C.
and Gendler,S.J.
Novel MUC1 splice variants contribute to mucin overexpression in
CPTX-deficient mice
Am. J. Physiol. Gastrointest. Liver Physiol. 284 (5), G853-G862
(2003)
JOURNAL MEDLINE 22570517
PUBMED 12529261
TITLE 2 (bases 1 to 1414)
Hinojosa-Kurtzberg,A.M. and Gendler,S.J.
Direct Submission
Submitted (21-SEP-2001) Biochemistry, Mayo Clinic, 13400 E. Shea
Boulevard, Scottsdale, AZ 85259, USA
JOURNAL location/Qualifiers
1. 1414
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/note="Homo sapiens gene in transgenic Mus musculus
C57BL/6; isolated from intestinal mucosa"
1. 1414
/gene="MUC1"
1. 1386
/gene="MUC1"
/note="alternatively spliced; contains exon 6b resulting
in variant carboxy-terminal domain; lacks sites for
beta-catenin and Grb2 interactions; derived from Homo
sapiens"
/codon_start=1
/transl_table=11
/product="mucin variant MUC1-CT58"
/protein_id="AA186735.1"
/db_xref="GI:19338622"
/translation="MTPGTQSPFFLLLLTLVTVVTSGHASSTPGGEKETATQSS
VPSSTERNVAVSMSTSVLSHSPGSGSSTTGGDVTIAPATEPASAATWGOVTSVP
VTRPALGSTPPAHVYTSAPDNKPAAGSTAPPAHGVTSAPDTPAAGVTS
APDNPALGSTAPVPHNVTSASGSASSTLVNCTSARATTPASKSTPSISHH
SDPTTLASHSTKTASSSTHSTVPPPLTSSNHTSPOLSTGVSAFFFLSHISNLOFNS
SLDPSITDYQELQDISSEMTQIYKGGFLGSLNFKRPGSVVVQTLAIFREGTINV

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 63.7%; Score 38.2; DB 6; Length 1774;
Best Local Similarity 78.0%; Pred. No. 0.56;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

OY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCAGCTCCTGACACAAGGCCAGCCCC 59
DB 462 GGCTCCACCGCCCCCGCCAGCCACGGTGTCACTCGGCCCGGACACCAAGCGCGCCCC 520

RESULT 9
AX959912 1774 bp DNA linear PAT 14-JAN-2004
LOCUS Sequence 17 from Patent WO03100060.
DEFINITION AX959912
ACCESSION AX959912
VERSION AX959912.1 GI:40880142
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 Burden,N.G., Ellis,J.H. and Hamblin,P.A.
AUTHORS Muc-1 antigen with reduced number of vntx repeat units
TITLE Patent: WO 03100060-A 17 04-DEC-2003;
JOURNAL GLAXO GROUP LIMITED (GB)
FEATURES location/Qualifiers
source 1. .1774
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 63.7%; Score 38.2; DB 6; Length 1774;
Best Local Similarity 78.0%; Pred. No. 0.56;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

OY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCAGCTCCTGACACAAGGCCAGCCCC 59
DB 462 GGCTCCACCGCCCCCGCCAGCCACGGTGTCACTCGGCCCGGACACCAAGCGCGCCCC 520

RESULT 10
AR492318 1804 bp DNA linear PAT 15-MAY-2004
LOCUS AR492318
DEFINITION Sequence 17 from patent US 6716627.
ACCESSION AR492318
VERSION AR492318.1 GI:47260892
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 Unclassified.
AUTHORS Dobie,K.W.
TITLE Antisense modulation of mucin 1, transmembrane expression
JOURNAL Patent: US 6716627-A 17 06-APR-2004;
FEATURES location/Qualifiers
source 1. .1804
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 63.7%; Score 38.2; DB 6; Length 1804;
Best Local Similarity 78.0%; Pred. No. 0.56;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

OY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCAGCTCCTGACACAAGGCCAGCCCC 59
DB 457 GGCTCCACCGCCCCCGCCAGCCACGGTGTCACTCGGCCCGGACACCAAGCGCGCCCC 515

RESULT 11

AX335367 1804 bp DNA linear PAT 09-JAN-2002
LOCUS AX335367
DEFINITION Sequence 5876 from Patent WO0194629.
ACCESSION AX335367
VERSION AX335367.1 GI:18126086
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
AUTHORS Horrigan,S., Soppet,D.R. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature
JOURNAL gene sets
PATENT: WO 0194629-A 5876 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES location/Qualifiers
source 1. .1804
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 63.7%; Score 38.2; DB 6; Length 1804;
Best Local Similarity 78.0%; Pred. No. 0.56;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

OY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCAGCTCCTGACACAAGGCCAGCCCC 59
DB 457 GGCTCCACCGCCCCCGCCAGCCACGGTGTCACTCGGCCCGGACACCAAGCGCGCCCC 515

RESULT 12
HUMMUCAB 1804 bp mRNA linear PRI 07-JAN-1995
LOCUS HUMMUCAB
DEFINITION Human polymorphic epithelial mucin (PEM) mRNA, complete cds.
ACCESSION J05581
VERSION J05581.1 GI:188869
KEYWORDS polymorphic epithelial mucin.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1804)
AUTHORS Gendler,S.J., Lancaster,C.A., Taylor-Papadimitriou,J., Duhig,T.,
TITLE Peat,N., Burchell,J., Pemberton,L., Lalani,E.N. and Wilson,D.
JOURNAL Molecular cloning and expression of human tumor-associated
MEDLINE polymorphic epithelial mucin
PUBMED J. Biol. Chem. 265 (25), 15286-15293 (1990)
COMMENT 1697589
Original source text: Homo sapiens adult adenocarcinoma cDNA to
mRNA.

FEATURES
source 1. .1804
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="Xq26.3-q27.1"
/cell_line="BT20"
/tissue_type="adenocarcinoma"
/dev_stage="adult"
1. .1804
/gene="MCNA"
/gene="MCNA"
/gene="MCNA"
/gene="MCNA"

gene
mRNA

```
/product="PEM mRNA"
73. .1500
/gene="MCNAA"
/codon_start=1
/product="polymorphic epithelial mucin"
/protein_id="AAAS9876.1"
/db_xref="GI:188870"
/translation="MTPGTQSPFFLLLLTVLTVTGSGHASSTPGKEKTSATORSS
VPSTERNKAVSMTSSVLSHSPGSGSTTQGDVTLAPATEPASGAATWQDVTSP
VTRPALGSTPPPAHDVTSAPDNKPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTS
APDNRPALGSTAPPVHNVTSASGSASGASTLVHNGTSARATTPPASKTFFSIPSHH
SDPTTLASHSTKTDASTHSTVPLTSSNHSSTPOLSTGVSEFFLSFHSINLQFNS
SLEDPSDYQELQORDISEMFLQIKQGGFLGLSNIKFRPGSVVQLTLAFREGTINV
HDVETQFNQYKTEASARYNLTISDVSVDVPEPSAOSGAGVPGMGIALVLVCLVA
LAIVYLIALAVCCQRRKNYGOLDIFPARDTYHPMSEYPTYHTHGRYVPPSSTDSPYE
KVSAGNGSGSLSYTNPAVAATSANL"
sig_peptide
73. .135
/gene="MCNAA"
/mat_peptide
136. .1497
/gene="MCNAA"
/product="polymorphic epithelial mucin"
1783. .1788
/gene="MCNAA"

polyA_signal

ORIGIN

Query Match      63.7%; Score 38.2; DB 9; Length 1804;
Best Local Similarity 78.0%; Pred. No. 0.56;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY      1 GGGTCGACTGCCCTCCGGCGCATGCTGTGACCTCCTGACACAAGCCAGCCCC 59
      |||||
Db      457 GGCTCCACCGCCCCCGACCGGTGTCACTCGGCCCCGGACACCAAGCCGCCCC 515

RESULT 13
AF423030      1834 bp      mRNA      linear      SYN 10-JUL-2003
LOCUS      Synthetic construct Homo sapiens mucin variant MUC-1CT80 (MUC1)
DEFINITION      mRNA, complete cds; alternatively spliced.
ACCESSION      AF423030
VERSION      AF423030.1 GI:19338619
KEYWORDS      .
SOURCE      synthetic construct
ORGANISM      synthetic construct
REFERENCE      1 (bases 1 to 1834)
AUTHORS      Hinojosa-Kurtzberg,A.M., Johansson,M.E., Madsen,C.S., Hansson,G.C.
and Gendler,S.J.
TITLE      Novel MUC1 splice variants contribute to mucin overexpression in
CFTR-deficient mice
JOURNAL      Am. J. Physiol. Gastrointest. Liver Physiol. 284 (5), G853-G862
(2003)
MEDLINE      22570517
PUBMED      12529261
REFERENCE      2 (bases 1 to 1834)
AUTHORS      Hinojosa-Kurtzberg,A.M. and Gendler,S.J.
TITLE      Direct Submision
JOURNAL      Submitted (21-SEP-2001) Biochemistry, Mayo Clinic, 13400 E. Shea
Boulevard, Scottsdale, AZ 85259, USA
FEATURES
source
1. .1834
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/note="Homo sapiens gene in transgenic Mus musculus
C57BL/6; isolated from intestinal mucosa"
1. .1834
/gene="MUC1"
1. .1452
/gene="MUC1"
/note="alternatively spliced; contains exon 6a resulting
in variant hydrophilic tail of carboxy-terminal domain;
lacks sites for beta-catenin and Grb2 interactions;
```

```
derived from Homo sapiens"
/codon_start=1
/translation="mucin variant MUC-1CT80"
/protein_id="AAL86734.1"
/db_xref="GI:19338620"
/translation="MTPGTQSPFFLLLLTVLTVTGSGHASSTPGKEKTSATORSS
VPSTERNKAVSMTSSVLSHSPGSGSTTQGDVTLAPATEPASGAATWQDVTSP
VTRPALGSTPPPAHDVTSAPDNKPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTS
APDNRPALGSTAPPVHNVTSASGSASGASTLVHNGTSARATTPPASKTFFSIPSHH
SDPTTLASHSTKTDASTHSTVPLTSSNHSSTPOLSTGVSEFFLSFHSINLQFNS
SLEDPSDYQELQORDISEMFLQIKQGGFLGLSNIKFRPGSVVQLTLAFREGTINV
HDVETQFNQYKTEASARYNLTISDVSVDVPEPSAOSGAGVPGMGIALVLVCLVA
LAIVYLIALAVCCQRRKNYGOLDIFPARDTYHPMSEYPTYHTHGRYVPPSSTDSPYE
KVLGPTQGKQRVWLGDSEGCTWKTQRAWKR"

ORIGIN

Query Match      63.7%; Score 38.2; DB 12; Length 1834;
Best Local Similarity 78.0%; Pred. No. 0.55;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY      1 GGGTCGACTGCCCTCCGGCGCATGCTGTGACCTCCTGACACAAGCCAGCCCC 59
      |||||
Db      385 GGCTCCACCGCCCCCGACCGGTGTCACTCGGCCCCGGACACCAAGCCGCCCC 443

RESULT 14
AX959918      1835 bp      DNA      linear      PAT 14-JAN-2004
LOCUS      Sequence 23 from Patent WO03100060.
DEFINITION      AX959918
ACCESSION      AX959918
VERSION      AX959918.1 GI:40880145
KEYWORDS      .
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      1
AUTHORS      Burden,N.G., Ellis,J.H. and Hamblin,P.A.
TITLE      Muc-1 antigen with reduced number of vnter repeat units
JOURNAL      Patent: WO 03100060-A 23 04-DEC-2003;
GLAXO GROUP LIMITED (GB)
FEATURES
source
1. .1835
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match      63.7%; Score 38.2; DB 6; Length 1835;
Best Local Similarity 78.0%; Pred. No. 0.55;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY      1 GGGTCGACTGCCCTCCGGCGCATGCTGTGACCTCCTGACACAAGCCAGCCCC 59
      |||||
Db      462 GGCTCCACCGCCCCCGACCGGTGTCACTCGGCCCCGGACACCAAGCCGCCCC 520

RESULT 15
AX959916      2135 bp      DNA      linear      PAT 14-JAN-2004
LOCUS      Sequence 21 from Patent WO03100060.
DEFINITION      AX959916
ACCESSION      AX959916
VERSION      AX959916.1 GI:40880144
KEYWORDS      .
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      1
AUTHORS      Burden,N.G., Ellis,J.H. and Hamblin,P.A.
TITLE      Muc-1 antigen with reduced number of vnter repeat units
JOURNAL      Patent: WO 03100060-A 21 04-DEC-2003;
```

FEATURES GLAXO GROUP LIMITED (GB)
Location/Qualifiers
Source 1. .2135
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 63.7%; Score 38.2; DB 6; Length 2135;
Best Local Similarity 78.0%; Pred. No. 0.53;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCAGCTCCTGACACAAAGGCCAGCCCC 59
Db 462 GGCTCCACCGCCCCCCCCCAGCCCAAGGTGTCACTCGCCCCCGACACCAAGCCGGCCCC 520

Search completed: January 15, 2005, 22:36:54
Job time : 723.8 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: January 15, 2005, 22:37:00 ; Search time 183.4 Seconds
(without alignments)
1879.789 Million cell updates/sec

Title: US-10-057-136-9

Perfect score: 60

Sequence: 1 ggaagtcacgcctccacactgc.....cagacactcgactcgcgcca 60

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq: *
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq: *
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq: *
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq: *
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq: *
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq: *
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq: *
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq: *
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq: *
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq: *
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq: *
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq: *
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq: *
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq: *
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq: *
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq: *
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq: *
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq: *
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq: *
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq: *
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	14 US-10-057-136-9	Sequence 9, Appli
2	60	100.0	2297	16 US-10-406-317-41	Sequence 41, Appl
3	52	86.7	60	14 US-10-057-136-8	Sequence 8, Appli
4	38.2	63.7	60	14 US-10-057-136-12	Sequence 12, Appl
5	33.4	55.7	60	14 US-10-057-136-4	Sequence 4, Appli
6	31.8	53.0	60	14 US-10-057-136-5	Sequence 5, Appli
7	31.4	52.3	93	17 US-10-471-607-3	Sequence 3, Appli
8	31.4	52.3	156	17 US-10-471-607-5	Sequence 5, Appli
9	31.4	52.3	157	17 US-10-471-607-6	Sequence 6, Appli
10	31.4	52.3	157	17 US-10-471-607-9	Sequence 9, Appli
11	31.4	52.3	3621	17 US-10-471-607-11	Sequence 11, Appl
12	31.4	52.3	6877	17 US-10-471-607-13	Sequence 13, Appl

13	30	50.0	60	14 US-10-057-136-6	Sequence 6, Appli
14	29	48.3	60	14 US-10-057-136-7	Sequence 7, Appli
15	29	48.3	60	14 US-10-057-136-11	Sequence 11, Appl
16	28	46.7	1845	15 US-10-156-761-3619	Sequence 3619, Ap
17	28	46.7	9025608	15 US-10-156-761-1	Sequence 1, Appli
18	27.4	45.7	60	14 US-10-057-136-2	Sequence 2, Appli
19	27.4	45.7	1424	16 US-10-447-839A-75	Sequence 75, Appl
20	27.4	45.7	1428	16 US-10-447-839A-20	Sequence 20, Appl
21	27.4	45.7	1527	14 US-10-057-136-19	Sequence 19, Appl
22	27.4	45.7	1799	16 US-10-447-839A-19	Sequence 19, Appl
23	27.4	45.7	1804	9 US-09-964-824A-573	Sequence 573, App
24	27.4	45.7	1804	15 US-10-029-517-17	Sequence 17, Appl
25	27.4	45.7	1804	17 US-10-717-597-30	Sequence 30, Appl
26	27.4	45.7	1804	17 US-10-775-920-84	Sequence 84, Appl
27	27.4	45.7	1823	15 US-10-101-510-339	Sequence 339, App
28	27.4	45.7	4139	9 US-09-964-824A-105	Sequence 105, App
29	27.4	45.7	4139	9 US-09-964-824A-578	Sequence 578, App
30	27.4	45.7	4139	9 US-09-864-864-334	Sequence 334, App
31	27.4	45.7	4139	9 US-09-880-107-2121	Sequence 2121, Ap
32	27.4	45.7	4139	11 US-09-968-007A-751	Sequence 751, App
33	27.4	45.7	4139	14 US-10-171-311-157	Sequence 157, App
34	27.4	45.7	4139	15 US-10-177-293-310	Sequence 310, App
35	27.4	45.7	4139	16 US-10-440-464-155	Sequence 155, App
36	27.4	45.7	4139	17 US-10-734-564-53	Sequence 53, Appl
37	27.4	45.7	4139	17 US-10-775-920-80	Sequence 80, Appl
38	27.4	45.7	4139	17 US-10-775-920-85	Sequence 85, Appl
39	27.4	45.7	8186	16 US-10-447-839A-18	Sequence 18, Appl
40	27.4	45.7	8186	14 US-10-247-703-23	Sequence 23, Appl
41	27.4	45.7	8186	15 US-10-029-517-19	Sequence 19, Appl
42	26.2	43.7	2649	15 US-10-369-493-42923	Sequence 42923, A
43	26	43.3	518	14 US-10-247-703-38	Sequence 38, Appl
44	26	43.3	518	15 US-10-029-517-101	Sequence 101, App
45	25.8	43.0	78	14 US-10-057-136-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-10-057-136-9
; Sequence 9, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-9

Query Match 100.0%; Score 60; DB 14; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.9e-12;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GGAAGTACCGCTCCACCTGCACACGGGGTACACAGCGCCAGACACTGCAGCTGCGCCA 60
|||||

Db

1 GGAAGTACCGCTCCACCTGCACACGCGGGTCAAGCGCGCCAGACACTCGACTGCGCCA 60

RESULT 2

US-10-406-317-41
; Sequence 41, Application US/10406317
; Publication No. US20040019195A1
; GENERAL INFORMATION:
; APPLICANT: Schiom, Jeffrey;
; APPLICANT: Hodge, James;
; APPLICANT: Panicali, Dennis
; TITLE OF INVENTION: A recombinant vector expressing multiple constitutulatory
; FILE REFERENCE: 38163-0189
; CURRENT APPLICATION NUMBER: US/10/406,317
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US/09/856,988
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: PCT/US99/26866
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/111,582
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 41
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: VECTOR
US-10-406-317-41

Query Match

Best Local Similarity 100.0%; Score 60; DB 16; Length 2297;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGCGGGTCAAGCGCGCCAGACACTCGACTGCGCCA 60

Db 706 GGAAGTACCGCTCCACCTGCACACGCGGGTCAAGCGCGCCAGACACTCGACTGCGCCA 765

RESULT 3

US-10-057-136-8
; Sequence 8, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHIOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 8
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-8

Query Match

Best Local Similarity 86.7%; Score 52; DB 14; Length 60;
Matches 91.7%; Pred. No. 1.6e-09;

Matches 55; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGCGGGTCAAGCGCGCCAGACACTCGACTGCGCCA 60

Db 1 GGCAGCACCGCACCGCGCCGACACGCGGTCAAGCGCGCCAGACACTCGACTGCGCCA 60

RESULT 4

US-10-057-136-12
; Sequence 12, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHIOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 12
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-12

Query Match

Best Local Similarity 63.7%; Score 38.2; DB 14; Length 60;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGCGGGTCAAGCGCGCCAGACACTCGACTGCGGCC 59

Db 1 GGTAGTACAGCGCCACCCGCACATGCGGTCAAGCGCGCTCCGATACAGACCGCGGCC 59

RESULT 5

US-10-057-136-4
; Sequence 4, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHIOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 4
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-4

Query Match	55.78;	Score 33.4;	DB 14;	Length 60;
Best Local Similarity	72.98;	Pred. No. 0.011;		
Matches 43;	Conservative	0;	Mismatches 16;	Indels 0;
				Gaps 0;

```

RESULT 6
US-10-057-136-5
; Sequence 5, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; TITLE OF INVENTION: TUMOR-ASSOCIATED ANTIGEN
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-5

```

```

RESULT 7
US-10-471-607-3
; Sequence 3, Application US/10471607
; Publication No. US20040115740A1
; GENERAL INFORMATION:
; APPLICANT: The Victoria University of Manchester
; APPLICANT: Benson, Roderick
; TITLE OF INVENTION: Intracellular analysis.
; FILE REFERENCE: P088657PWO
; CURRENT APPLICATION NUMBER: US/10/471,607
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: GB 0108165.2
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 93
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Artificial epitope construct
US-10-471-607-3

Query Match          52.3%; Score 31.4; DB 17; Length 93;
Best Local Similarity 85.4%; Pred.No. 0.058;

```

```

Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 GGAAGTACCGCTCCACCTGCACACGGGGTCAAGCGGCC 41
    ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 38 GGATCTACCGCTCTCTGCCCCAGGAGTCACAAGCGCACC 78

RESULT 8
US-10-471-607-5
; Sequence 5, Application US/10471607
; Publication No. US20040115740A1
; GENERAL INFORMATION:
; APPLICANT: The Victoria University of Manchester
; APPLICANT: Benson, Roderick
; TITLE OF INVENTION: Intracellular analysis.
; FILE REFERENCE: P088857PMO
; CURRENT APPLICATION NUMBER: US/10/471,607
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: GB 0108165.2
; PRIOR FILING DATE: 3001-03-21
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 156
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Artificial epitope construct
US-10-471-607-5

```

```

RESULT 9
US-10-471-607-6
; Sequence 6, Application US/10471607
; Publication No. US20040115740A1
; GENERAL INFORMATION:
; APPLICANT: The Victoria University of Manchester
; APPLICANT: Benson, Roderick
; TITLE OF INVENTION: Intracellular analysis.
; FILE REFERENCE: P08857PMO
; CURRENT APPLICATION NUMBER: US/10/471,607
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: GB 0108165.2
; PRIOR FILING DATE: 3001-03-21
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 157
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Artificial epitope construct
US-10-471-607-6

Query Match          52.3%; Score 31.4; DB 17; Length 157;
Best Local Similarity 85.4%; Pred. No. 0.058;
Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      1  GGAAGTACCGCTCCACCTGCACACAGGGGTCAACAAGCGCGCC 41
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      38  GGATCTACCGCTCTCTGCGCCACGAGATCAACAAGCGCAC 78

```



```
; Sequence 9, Application US/10471607
; Publication No. US20040115740A1
; GENERAL INFORMATION:
; APPLICANT: The Victoria University of Manchester
; APPLICANT: Benson, Roderick
; TITLE OF INVENTION: Intracellular analysis.
; FILE REFERENCE: P088857PMO
; CURRENT APPLICATION NUMBER: US/10/471,607
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: GB 0108165.2
; PRIOR FILING DATE: 3001-03-21
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 157
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Artificial epitope construct
US-10-471-607-9
```

```
Query Match
Best Local Similarity 52.3%; Score 31.4; DB 17; Length 157;
Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
Qy 1 GGAAGTACCGCTCCACCTGCACACGGGGTCACAAGCGCGCC 41
Db 38 GGATCTACCGCTCTCTGCCCCACGGAGTCACAAGCGCGACC 78
```

```
RESULT 11
US-10-471-607-11
; Sequence 11, Application US/10471607
; Publication No. US20040115740A1
; GENERAL INFORMATION:
; APPLICANT: The Victoria University of Manchester
; APPLICANT: Benson, Roderick
; TITLE OF INVENTION: Intracellular analysis.
; FILE REFERENCE: P088857PMO
; CURRENT APPLICATION NUMBER: US/10/471,607
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: GB 0108165.2
; PRIOR FILING DATE: 3001-03-21
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 3621
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Engineered construct
US-10-471-607-11
```

```
Query Match
Best Local Similarity 52.3%; Score 31.4; DB 17; Length 3621;
Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
Qy 1 GGAAGTACCGCTCCACCTGCACACGGGGTCACAAGCGCGCC 41
Db 659 GGATCTACCGCTCTCTGCCCCACGGAGTCACAAGCGCGACC 699
```

```
RESULT 12
US-10-471-607-13
; Sequence 13, Application US/10471607
; Publication No. US20040115740A1
; GENERAL INFORMATION:
; APPLICANT: The Victoria University of Manchester
; APPLICANT: Benson, Roderick
; TITLE OF INVENTION: Intracellular analysis.
; FILE REFERENCE: P088857PMO
; CURRENT APPLICATION NUMBER: US/10/471,607
; CURRENT FILING DATE: 2003-09-24
```

```
; PRIOR APPLICATION NUMBER: GB 0108165.2
; PRIOR FILING DATE: 3001-03-21
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 6877
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Engineered construct
US-10-471-607-13
```

```
Query Match
Best Local Similarity 52.3%; Score 31.4; DB 17; Length 6877;
Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
Qy 1 GGAAGTACCGCTCCACCTGCACACGGGGTCACAAGCGCGCC 41
Db 701 GGATCTACCGCTCTCTGCCCCACGGAGTCACAAGCGCGACC 741
```

```
RESULT 13
US-10-057-136-6
; Sequence 6, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-6
```

```
Query Match
Best Local Similarity 50.0%; Score 30; DB 14; Length 60;
Matches 39; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
```

```
Qy 7 ACCGCTCACCCTGCACACGGGGTCACAAGCGGCCAGACACTGCACCTGGCGCCA 60
Db 7 ACAGCTCTCTCCCGCTCATGGGTTACTTCTGCTCCAGATACCTGCCCGACTCCA 60
```

```
RESULT 14
US-10-057-136-7
; Sequence 7, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
```

```
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-7
```

```
Query Match          48.3%; Score 29; DB 14; Length 60;
Best Local Similarity 71.7%; Pred. No. 0.45;
Matches 38; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
```

```
QY      7 ACCGCTCCACCTGCACACGGGGTCAAGCGCCGACACTCGACTGCGCC 59
      |||||
Db      7 ACGGCCCCCCTGCTCAGGTGTACATCCGCCCCGGATACCAAGACCGCCCC 59
```

```
RESULT 15
US-10-057-136-11
```

```
; Sequence 11, Application US/10057136
; Publication No. US2003002170A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFU, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-11
```

```
Query Match          48.3%; Score 29; DB 14; Length 60;
Best Local Similarity 71.7%; Pred. No. 0.45;
Matches 38; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
```

```
QY      7 ACCGCTCCACCTGCACACGGGGTCAAGCGCCGACACTCGACTGCGCC 59
      |||||
Db      7 ACGGCACTTCACACAGGAGTACGTCGACCCGACACCCGTCAGCTCC 59
```

Search completed: January 16, 2005, 09:30:30
Job time : 188.4 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 19:43:55 ; Search time 1550.6 Seconds
(without alignments)
1410.024 Million cell updates/sec

Title: US-10-057-136-9
Perfect score: 60
Sequence: 1 ggaagtaccgcctccacactgc.....cagacactcgacctgcgcga 60

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29.8	49.7	653	4 BM582925	BM582925 170006872
C 2	28.6	47.7	657	6 BY729679	BY729679 BY729679
C 3	28.2	47.0	513	9 CE582697	CE582697 tigr-gss-
C 4	28.2	47.0	514	9 CE347100	CE347100 tigr-gss-
5	28.2	47.0	603	9 CE362548	CE362548 tigr-gss-
6	28.2	47.0	652	5 BX611440	BX611440 BX611440
7	28.2	47.0	682	4 BM607934	BM607934 170006870
C 8	27.8	46.3	583	4 BG005521	BG005521 MR3-GN018
C 9	27.4	45.7	330	1 AI925867	AI925867 wo20d04.x
10	27.4	45.7	604	4 BM791359	BM791359 K-EST0071
11	27.4	45.7	877	5 BU542454	BU542454 AGENCOURT
12	27.4	45.7	890	5 CB994654	CB994654 AGENCOURT
13	27.4	45.7	959	6 CA489836	CA489836 AGENCOURT
14	27.4	45.7	1113	5 BU148487	BU148487 AGENCOURT
15	27.4	45.7	1130	5 BU542996	BU542996 AGENCOURT
16	27.4	45.7	1234	5 BQ936898	BQ936898 AGENCOURT
17	27.4	45.7	1262	5 BQ935496	BQ935496 AGENCOURT
18	27.4	45.7	1268	5 BQ943554	BQ943554 AGENCOURT
19	27.4	45.7	1343	5 BQ920055	BQ920055 AGENCOURT
20	27.4	45.7	1349	5 BU152566	BU152566 AGENCOURT
21	27.4	45.7	1420	5 BU542790	BU542790 AGENCOURT
22	27.4	45.7	1531	5 BU543309	BU543309 AGENCOURT
C 23	27	45.0	184	6 CA780008	CA780008 MPL384 3
24	27	45.0	393	5 BU653553	BU653553 1112107D0

25	27	45.0	394	4 BI722431	BI722431 1031062A0
C 26	27	45.0	447	1 AJ665988	AJ665988 AJ665988
C 27	27	45.0	501	5 BP436805	BP436805 BP436805
C 28	27	45.0	521	5 BP152756	BP152756 BP152756
C 29	27	45.0	572	5 BP455789	BP455789 BP455789
C 30	27	45.0	573	5 BP148041	BP148041 BP148041
C 31	27	45.0	585	5 BP169164	BP169164 BP169164
C 32	27	45.0	602	5 BP148340	BP148340 BP148340
C 33	27	45.0	628	5 BP147720	BP147720 BP147720
C 34	27	45.0	648	5 BP152406	BP152406 BP152406
C 35	27	45.0	655	5 BP152491	BP152491 BP152491
C 36	27	45.0	663	5 BP461462	BP461462 BP461462
C 37	27	45.0	672	5 BP151509	BP151509 BP151509
C 38	27	45.0	675	5 BP148090	BP148090 BP148090
C 39	27	45.0	676	5 BP151085	BP151085 BP151085
C 40	27	45.0	676	5 BP153615	BP153615 BP153615
C 41	27	45.0	683	5 BP151728	BP151728 BP151728
C 42	27	45.0	684	5 BP435328	BP435328 BP435328
C 43	27	45.0	692	5 BP150259	BP150259 BP150259
C 44	27	45.0	697	5 BP455277	BP455277 BP455277
C 45	27	45.0	705	5 BP146979	BP146979 BP146979

ALIGNMENTS

RESULT 1
LOCUS BM582925 653 bp mRNA linear EST 22-FEB-2002
DEFINITION 17000687277260 A.Gam.ad.cDNA.blood1 Anopheles gambiae cDNA clone
19600449716094 5', mRNA sequence.
ACCESSION BM582925
VERSION BM582925.1 GI:18871392
KEYWORDS EST.
SOURCE Anopheles gambiae (African malaria mosquito)
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
Anopheles.
REFERENCE 1 (bases 1 to 653)
AUTHORS Holt,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L.,
Charlab,R., Collins,F.H., Venter,J.C. and Hoffman,S.L.
TITLE Celera Anopheles gambiae EST project
JOURNAL Unpublished (2002)
COMMENT Contact: Holt R.A.
Celera Genomics
45 W. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151
Fax: 2404534580
Email: HoltRA@celera.com
Plate: NU01004HLX row: G column: 24
Seq primer: M13 Reverse.
FEATURES
source location/Qualifiers
1..653
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="RSP-ST (Reduced succ. to Permethrin - std.
chromosome)"
/db_xref="taxon:7165"
/clone="19600449716094"
/dev_stage="Adult"
/lab_host="DH10b"
/clone_lib="A.Gam.ad.cDNA.blood1"
/note="Vector: pSport1; Site 1: SalI; Site 2: NotI; whole
adult mosquitoes (mixed sex) frozen on liquid nitrogen 24
hours after human blood feeding. cDNA inserts >500 bp
cloned directionally into pSport 1. Not 1 site is 3'.
Clones available through the Malaria Research and
Reference Reagent Resource Center (www.malaria.mr4.org)"
ORIGIN
Query Match 49.7%; Score 29.8; DB 4; Length 653;
Best Local Similarity 70.2%; Pred. No. 20;

Matches 40; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACAGCGGGTCACAGCGCGCCAGACACTGCACCTGCG 57
Db 24 GGAGGGTCCCTTATCTGTGCTCACCATGTCACTAGCGTGCACACACCCGACCTGTG 80

RESULT 2
BY729679/c

LOCUS
DEFINITION BY729679 657 bp mRNA linear EST 17-DEC-2002
musculus CDNA clone C530042M15 5', mRNA sequence.
BY729679
BY729679.1 GI:27142806

EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 657)

REFERENCE
AUTHORS
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gotohori, T., Baldarelli, R., Hill, D.P., Bult, C.,
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,
Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
Gariboldi, M., Glisai, C., Godzik, A., Gough, J., Grimmond, S.,
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
Sandelin, A., Schneider, C., Sempile, C.A., Setou, M., Shimada, K.,
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
22354683
12466851

JOURNAL MEDLINE PUBMED

TITLE
COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P.,
Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F.,
Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y.,
Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M.,
Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N.,
Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,
Takeda, Y., Waki, K., Watanishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN,
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.

FEATURES
source
Location/Qualifiers

1..657
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="C530042M15"
/tissue_type="spinal cord"
/dev_stage="12 days embryo"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 12 days embryo
spinal_cord"
/note="Site 1: SalI; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTCTTTVN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5',
GAGAGAGAGATTCGAGTTAATTAATTAATCCCCCCCCCC 3']. cDNA
was cleaved with XhoI and BamHI. Vector: a modified
pBluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN

Query Match 47.7%; Score 28.6; DB 6; Length 657;
Best Local Similarity 67.8%; Pred. No. 48;
Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACAGCGGGTCACAGCGCGCCAGACACTGCACCTGCGC 59
Db 485 GGAGGGTCCCTTATCTGTGCTCACCATGTCACTAGCGTGCACACACCCGACCTGTG 80

RESULT 3
CE582697/c

LOCUS
DEFINITION tigr-gss-dog-17000366290475 Dog Library Canis familiaris genomic,
genomic survey sequence.
CE582697
CE582697.1 GI:36899478

VERSION
KEYWORDS
SOURCE
ORGANISM
Canis familiaris (dog)
Canis familiaris

REFERENCE
AUTHORS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Placentalia; Canidae; Canis.
1 (bases 1 to 513)
Kirkness, E.F., Batina, V., Halpern, A.L., Levy, S., Remington, K.,
Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and
Venter, J.C.

TITLE
JOURNAL MEDLINE PUBMED
The dog genome: survey sequencing and comparative analysis
Science 301 (5641), 1898-1903 (2003)
22875432
14512627

COMMENT
Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,

Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.

FEATURES
Source
Location/Qualifiers
1..513

/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: Batxi; Libraries were prepared from
peripheral blood"

ORIGIN

Query Match 47.0%; Score 28.2; DB 9; Length 513;
Best Local Similarity 73.5%; Pred. No. 65;
Matches 36; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 8 CCGCTCCACCTGCACACGGGGTCAACAAGCGCCAGACTCGACCTGC 56
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 424 CCCCAGCACTGCATCGGGGTCCCAAGTGCAGAGCCGCACCTGC 376

RESULT 4
CE347100/c 514 bp DNA linear GSS 26-SEP-2003
LOCUS tigr-gss-dog-1700034141252 Dog Library Canis familiaris genomic,
DEFINITION genomic survey sequence.
ACCESSION CE347100
VERSION CE347100.1 GI:36177454
KEYWORDS GSS.

SOURCE
ORGANISM Canis familiaris (dog)
Canis familiaris

REFERENCE
AUTHORS Kirkenes, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,
Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and
Venter, J.C.

TITLE The dog genome: survey sequencing and comparative analysis
JOURNAL Science 301 (5641), 1898-1903 (2003)
MEDLINE 22875432
PUBMED 14512627

COMMENT
Contact: Kirkness EF

The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org

Class: shotgun.

FEATURES
Source
Location/Qualifiers
1..514

/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: Batxi; Libraries were prepared from
peripheral blood"

ORIGIN

Query Match 47.0%; Score 28.2; DB 9; Length 514;
Best Local Similarity 73.5%; Pred. No. 65;
Matches 36; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 8 CCGCTCCACCTGCACACGGGGTCAACAAGCGCCAGACTCGACCTGC 56
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 51 CCCCAGCACTGCATCGGGGTCCCAAGTGCAGAGCCGCACCTGC 3

RESULT 5
CE362548 603 bp DNA linear GSS 27-SEP-2003
LOCUS tigr-gss-dog-17000361549348 Dog Library Canis familiaris genomic,
DEFINITION genomic survey sequence.

ACCESSION CE362548
VERSION CE362548
KEYWORDS GSS.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris

REFERENCE
AUTHORS Kirkenes, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,
Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and
Venter, J.C.

The dog genome: survey sequencing and comparative analysis
Science 301 (5641), 1898-1903 (2003)

JOURNAL 22875432
MEDLINE 14512627
PUBMED

COMMENT
Contact: Kirkness EF

The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org

Class: shotgun.

FEATURES
Source
Location/Qualifiers
1..603

/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: Batxi; Libraries were prepared from
peripheral blood"

ORIGIN

Query Match 47.0%; Score 28.2; DB 9; Length 603;
Best Local Similarity 73.5%; Pred. No. 65;
Matches 36; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 8 CCGCTCCACCTGCACACGGGGTCAACAAGCGCCAGACTCGACCTGC 56
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 141 CCCCAGCACTGCATCGGGGTCCCAAGTGCAGAGCCGCACCTGC 189

RESULT 6

BX611440 652 bp mRNA linear EST 07-AUG-2003
LOCUS BX611440 Normalized Anopheles Head (NAH) Library Anopheles gambiae
DEFINITION cDNA clone AGA781TR, mRNA sequence.

ACCESSION BX611440
VERSION BX611440.1 GI:33501327
KEYWORDS EST.
SOURCE Anopheles gambiae (African malaria mosquito)
ORGANISM Anopheles gambiae

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
Anopheles.

REFERENCE
1 (bases 1 to 652)

Lobo, N.L., Gardner, M., Romans, P. and Collins, F.H.
Anopheles gambiae EST, Center for Tropical Disease Research and
Training

JOURNAL Unpublished (2003)
AUTHORS Contact: Frank H. Collins
TITLE Center for Tropical Disease Research and Training
COMMENT University of Notre Dame

Notre Dame, IN 46556, USA
Tel: 574-631-9245
Fax: 574-631-3996
Email: frank.h.collins.75@nd.edu.

FEATURES

source

Location/Qualifiers

```
1..652
/organism="Anopheles gambiae"
/mol_type="mRNA"
/db_xref="taxon:7165"
/clone="AGAA781TR"
/lab_host="E. coli DH10B"
/clone_lib="Normalized Anopheles Head (NAH) Library"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI (5'end); Site_2: NotI (3'end); a
directionally cloned and normalized, oligo-T primed cDNA
library constructed from strain 4arr adult mosquito heads.
Equal numbers of sugar fed males, sugar fed females and 6,
24 and 48 hr post blood meal females were used: Bonaldo,
Lennon & Soares (1996): Normalization and Subtraction: Two
Approaches to Facilitate Gene Discovery, Genome Research
6, 791-806. ESTs sequenced from the M13 reverse priming
site reading from the 5' ends of the cDNAs are indicated
by 'R' in the clone name. ESTs sequenced from the M13
forward priming site reading from the 3' ends of the cDNAs
are indicated by 'F' in the clone name."
```

ORIGIN

Query Match 47.0%; Score 28.2; DB 5; Length 652;
Best Local Similarity 68.4%; Pred. No. 66;
Matches 39; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGCGGTCAACAGCGCGGCACTGCGCTGCG 57
DB 395 GGAGGGTCCCTCTATCTGCTCACCATGTCTACCTGCTGCACACGCCGACCTGTG 451

RESULT 7

LOCUS

BM607934 682 bp mRNA linear EST 25-FEB-2002
DEFINITION 17000687085787 A.Gam.ad.cDNA.blood1 Anopheles gambiae cDNA clone

19600449719648 5', mRNA sequence.

ACCESSION

BM607934

VERSION

BM607934.1 GI:18906038

KEYWORDS

EST.

SOURCE

Anopheles gambiae (African malaria mosquito)

ORGANISM

Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
Anopheles.

1 (bases 1 to 682)

REFERENCE

Holt,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L.,
Charlab,R., Collins,F.H., Venter,J.C. and Hoffman,S.L.,
Celera Anopheles gambiae EST project
Unpublished (2002)

JOURNAL

Contact: Holt R.A.

COMMENT

FEATURES

source

45 W. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151
Fax: 2404534580
Email: HoltRA@celera.com
Plate: NT010049VO row: L column: 02
Seq primer: M13 Reverse.

FEATURES

source

Location/Qualifiers

```
1..682
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="RSP-ST (Reduced susc. to Permethrin - std.
chromosome)"
/db_xref="taxon:7165"
/clone="19600449719648"
/dev_stage="Adult"
/lab_host="DH10B"
/clone_lib="A.Gam.ad.cDNA.blood1"
/note="Vector: pSport1; Site_1: SalI; Site_2: NotI; Whole
adult mosquitoes (mixed sex)-frozen on liquid nitrogen 24
hours after human blood feeding. cDNA inserts >500 bp
cloned directionally into pSport 1. Not 1 site is 3'.
```

Clones available through the Malaria Research and
Reference Reagent Resource Center (www.malaria.mr4.org) "

ORIGIN

Query Match 47.0%; Score 28.2; DB 4; Length 682;
Best Local Similarity 68.4%; Pred. No. 66;
Matches 39; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGCGGTCAACAGCGCGGCACTGCGCTGCG 57
DB 336 GGAGGGTCCCTCTATCTGCTCACCATGTCTACCTGCTGCACACGCCGACCTGTG 392

RESULT 8

LOCUS

BG005521/c

DEFINITION

MR3-GN0186-241100-008-b11 GN0186 Homo sapiens cDNA, mRNA sequence.

ACCESSION

BG005521

VERSION

BG005521.1 GI:12447774

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 583)

TITLE

Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

JOURNAL

PROC. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

20202663

PUBMED

10737800

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR3&t2=MR3-GN0186-241100-008-b11&t3=2000-11-24&t4=1>)

Seq primer: puc 18 forward

High quality sequence stop: 539.

Location/Qualifiers

FEATURES

source

1..583
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="GN0186"

/note="Organ: placenta_normal; Vector: puc18; Site_1:
SmaI; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the pUC 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

ORIGIN

Query Match 46.3%; Score 27.8; DB 4; Length 583;
Best Local Similarity 74.5%; Pred. No. 88;
Matches 35; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 13 CCACCTGCACACGGGTCAACAGCGCGGCACTGCGCTGCGCC 59
DB 83 CCACCTGCACACATTCCACACGCGGACACACACACTGCGACAGCGCC 37

RESULT 9
AI925867/c 330 bp mRNA linear EST 08-MAR-2000
LOCUS
DEFINITION wo20d04.x1 NCI CGAP Paul Homo sapiens cDNA clone IMAGE:2455879 3', similar to contains_element MSRI MSRI repetitive element ;, mRNA sequence.
ACCESSION AI925867
VERSION AI925867.1 GI:5661831
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
1 (bases 1 to 330)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bdrp/image/image.html
Insert length: 578 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 305.
Location/Qualifiers
source
1..330
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2455879"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/clone_lib="NCI CGAP Paul"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: Salt; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.72 kb. Life Technologies catalog #: 11548-013"
ORIGIN
Query Match 45.7%; Score 27.4; DB 1; Length 330;
Best Local Similarity 69.8%; Pred. No. 1.2e+02;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 7 ACCGCTCACCTGCACACGGGGTACACAGCGCGCACACTCGACTGCGCC 59
LOCUS |||||
DEFINITION 310 ACCGCCGCCCGACGCCGCTGTCACCTCGGCCCGGACACCGCGGCCCC 258
Db
RESULT 10
BM791359 604 bp mRNA linear EST 05-MAR-2002
LOCUS BM791359
DEFINITION K-EST0071342 S21SNUS20 Homo sapiens cDNA clone S21SNUS20-14-A06 5', mRNA sequence.
ACCESSION BM791359
VERSION BM791359.1 GI:19139591
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
1 (bases 1 to 604)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.
21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS

Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 14 row: A column: 06
High quality sequence stop: 604.
Location/Qualifiers
source
1..604
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S21SNUS20-14-A06"
/sex="F"
/tissue_type="Stomach"
/cell_type="Floating aggregates"
/cell_line="SNU-520"
/lab_host="Top10F"
/clone_lib="S21SNUS20"
/note="Organ: Stomach; Vector: pTZ19RP1; Site_1: EcoRI; Site_2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."
ORIGIN
Query Match 45.7%; Score 27.4; DB 4; Length 604;
Best Local Similarity 69.8%; Pred. No. 1.2e+02;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 7 ACCGCTCACCTGCACACGGGGTACACAGCGCGCACACTCGACTGCGCC 59
LOCUS |||||
DEFINITION 34 ACCGACCCCGACGCCGCTGTCACCTCGGCCCGGACACCGCGGCCCC 86
Db
RESULT 11
BUS42454 877 bp mRNA linear EST 13-SEP-2002
LOCUS BUS42454
DEFINITION AGENCOURT_10322173 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574322 5', mRNA sequence.
ACCESSION BUS42454
VERSION BUS42454.1 GI:22852937
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
1 (bases 1 to 877)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLCM2769 row: h column: 02

FEATURES High quality sequence stop: 760.
Location/Qualifiers
source 1..877
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6574322"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAAGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 45.7%; Score 27.4; DB 5; Length 877;
Best Local Similarity 69.8%; Pred. No. 1.2e+02;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACCGCTCCACCTGCACACGCGGGTCACAGCGCGCCAGACACTCGACCTGCGCC 59
Db 41 ACCGCCCCCCAGCCACCGGTTCACCTCGGCCCGGACCCAGGCCGCCCC 93

RESULT 12
CB994654 890 bp mRNA linear EST 01-MAY-2003
LOCUS
DEFINITION AGENCOURT_13627705 NIH_MGC_148 Homo sapiens cDNA clone
IMAGE:30335618 5', mRNA sequence.
ACCESSION CB994654
VERSION CB994654.1 GI:30289174
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
REFERENCE 1 (bases 1 to 890)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: NDAM357 row: m column: 03
High quality sequence stop: 503.
Location/Qualifiers
1..890
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30335618"
/tissue_type="pre-ecampic placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_148"
/note="Organ: placenta; Vector: pBluescriptR; Site 1:
all-xhoI; Site 2: BamH; Library is oligo-dT primed and
directionally cloned using primer
5'-TTTTTTTCTTTTCTTCA-3', size-selected for average insert
size 2.3 kb and normalized to ROT 5. This is a primary
library enriched for full-length clones and constructed
using the Cap-trapper method (Carninci, in preparation)."

FEATURES
source

Library constructed by M. Brownstein (NHGRI/NHGR),
National Institutes of Health). Note: this is a NIH_MGC
Library."

ORIGIN

Query Match 45.7%; Score 27.4; DB 6; Length 890;
Best Local Similarity 75.6%; Pred. No. 1.2e+02;
Matches 34; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGCGGGTCACAGCGCGCCAGAC 45
Db 721 GGTTGTACCGATCCACCTCCCGCAGAGTCTCAAGGCCCGCCAGAC 765

RESULT 13
CA489836 959 bp mRNA linear EST 14-NOV-2002
LOCUS
DEFINITION AGENCOURT_10810668 MAPcL Homo sapiens cDNA clone IMAGE:6722324 5',
mRNA sequence.
ACCESSION CA489836
VERSION CA489836.1 GI:24952627
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
REFERENCE 1 (bases 1 to 959)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Kristi A. Eglund, Ira Pastan
cDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM14284 row: n column: 20
High quality sequence start: 31
High quality sequence stop: 446.
Location/Qualifiers
1..959
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6722324"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231,
HRT-HEM1, LNCAP"
/lab_host="EMDH10B"
/clone_lib="MAPcL"
/note="Vector: pCMV-SPORT6; Site 1: EcoRV; Site 2: Not I;
Subtracted with brain, liver, lung, kidney and muscle.
Directionally cloned. Priming method: oligo-dT. Average
insert size: 1800 bp. Library amplification: 26,000 fold.
Kristi A. Eglund, James J. Vincent, Robert Strausberg,
Bungkook Lee & Ira Pastan: Discovery of new breast
cancer genes encoding membrane and secreted proteins.
Manuscript submitted."

ORIGIN

Query Match 45.7%; Score 27.4; DB 6; Length 959;
Best Local Similarity 69.8%; Pred. No. 1.2e+02;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACCGCTCCACCTGCACACGCGGGTCACAGCGCGCCAGACACTCGACCTGCGCC 59
Db 355 ACCGCCCCCCAGCCACCGGTTCACCTCGGCCCGGACACCAAGCGCGCCCC 407

RESULT 14
BU148487

LOCUS	1113 bp	mRNA	linear	EST 03-SEP-2002
DEFINITION	AGENCOURT_8670479 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6380642			
ACCESSION	5', mRNA sequence.			
VERSION	BU148487			
KEYWORDS	BU148487.1	GI:22662019		
SOURCE	EST.			
ORGANISM	Homo sapiens (human)			
	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 113)			
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .			
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL	Unpublished (1999)			
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: DCTD/DTP			
	cDNA Library Preparation: Rubin Laboratory			
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)			
	DNA Sequencing by: Agencourt Bioscience Corporation			
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov			
	plate: LLCM2569	row: j	column: 03	
	High quality sequence scop: 235.			
FEATURES	Location/Qualifiers			

ORIGIN

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6380642"
/issue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/notes="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

```

ORIGIN

Query Match	45.7%;	Score 27.4;	DB 5;	Length 1113;
Best Local Similarity	69.8%;	Pred. No. 1.2e+02;		
Matches 37;	Conservative	0;	Mismatches 16;	Indels 0;
				Gaps 0;

Oy 7 ACCGCTCCACCTGCACACAGGGGGTCAACAAGCGCGCCAGACACTTGACCTGCGCC 59
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 36 ACCGCCCCCACCAGCCCACAGGTGTTCACCTCTGCCCCGGACACCAAGCCCGGCC 88

RESULT 15
BUS42996

LOCUS	BU542996	1130 bp	mRNA	linear	EST 13-SEP-2002
-------	----------	---------	------	--------	-----------------

5', mRNA sequence. - - -

VERSION BU542996.1 GI:22853479

KEYWORDS EST.

SOURCE	Homo sapiens (human)
--------	----------------------

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Placentalia; Primates; Carnivora; Canidae; Felidae; Felis

1 (bases 1 to 1120)
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE
1 (DABBS 1 CO 1130)

TITLE National Institutes of Health. Mammalian Gene Collection (MGC)

JOURNAL
Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: DCTD/DTP

CDNA Library Preparation: Rubin Laboratory

```

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2771 row: c column: 11
High quality sequence start: 27
High quality sequence stop: 246.
    location/Qualifiers
1. .1130
FEATURES
Source
```

FEATURES

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6574979"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

```

ORIGIN

Query Match	45.7%	Score 27.4;	DB 5;	Length 1130;
Best Local Similarity	69.8%	Pred. No. 1.2e+02;		
Matches 37;	Conservative 0;	Mismatches 16;	Indels 0;	Gaps 0;

Qy	7	ACCGCTCACTGCACACGGGGTCA	CGCGCCAGACACTGCCTGCGCC	59
Db	75	ACGGCCCCCAGCCACCGGTGTCA	CTCGGCCCCGAGACACGAGCCGCCCC	127

Search completed: January 16, 2005, 02:55:36
Job time : 1553.6 secs

Job time : 1553.6 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 19:50:13 ; Search time 35.2 Seconds
(without alignments)
1211.572 Million cell updates/sec

Title: US-10-057-136-9

Perfect score: 60
Sequence: 1 ggaagtagcgcctccacactgc.....cagacactcgactcgcgcca 60

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27.4	45.7	1804	4 US-10-029-517-17	Sequence 17, Appl
2	27.4	45.7	8186	4 US-10-029-517-19	Sequence 19, Appl
3	26	43.3	518	4 US-10-029-517-101	Sequence 101, App
4	25.8	43.0	572	4 US-10-029-517-18	Sequence 18, Appl
5	25.8	43.0	1721	4 US-10-029-517-3	Sequence 3, Appl
6	24.6	41.0	60	4 US-09-475-947A-246	Sequence 246, App
7	24.4	40.7	846	4 US-09-455-960-5	Sequence 5, Appl
8	24.4	40.7	846	4 US-10-051-325-5	Sequence 5, Appl
9	24.4	40.7	2500	3 US-09-455-960-2	Sequence 2, Appl
10	24.4	40.7	2500	4 US-10-051-325-2	Sequence 2, Appl
11	24.4	40.7	3343	4 US-10-029-517-102	Sequence 102, App
12	23.6	39.3	1392	4 US-09-489-039A-5663	Sequence 5663, Ap
13	23.6	39.3	1772	3 US-09-276-531-86	Sequence 86, Appl
14	23.2	38.7	507	4 US-08-956-171E-3895	Sequence 3895, Ap
15	23.2	38.7	507	4 US-08-781-986A-3895	Sequence 3895, Ap
16	23.2	38.7	558	1 US-08-729-767-1	Sequence 1, Appl
17	23.2	38.7	558	1 US-08-729-767-1	Sequence 1, Appl
18	23.2	38.7	657	4 US-09-252-991A-15827	Sequence 15827, A
19	23.2	38.7	2687	4 US-10-037-616-2	Sequence 2, Appl
20	23.2	38.7	3280	1 US-08-259-000-4	Sequence 4, Appl
21	23.2	38.7	3280	1 US-08-729-767-6	Sequence 6, Appl
22	23.2	38.7	3543	4 US-09-252-991A-15893	Sequence 15893, A
23	23.2	38.7	4257	4 US-09-252-991A-15796	Sequence 15796, A
24	23.2	38.7	6192	2 US-08-479-537A-1	Sequence 1, Appl
25	23.2	38.7	6192	3 US-09-083-116-1	Sequence 1, Appl
26	23.2	38.7	6192	3 US-09-134-916A-1	Sequence 1, Appl
27	23.2	38.7	6449	2 US-08-479-537A-4	Sequence 4, Appl

28	23.2	38.7	6449	3 US-09-083-116-4	Sequence 4, Appl
29	23.2	38.7	6449	3 US-09-134-916A-4	Sequence 4, Appl
30	23.2	38.7	13794	4 US-08-956-171E-54	Sequence 54, Appl
31	23.2	38.7	13794	4 US-08-781-986A-54	Sequence 54, Appl
32	23	38.3	461	4 US-09-621-976-3353	Sequence 3353, Ap
33	23	38.3	2680	4 US-09-566-921-9	Sequence 9, Appl
34	23	38.3	2727	4 US-09-919-497-13	Sequence 13, Appl
35	22.8	38.0	425	4 US-08-956-171E-3693	Sequence 3693, Ap
36	22.8	38.0	425	4 US-08-781-986A-3693	Sequence 3693, Ap
37	22.8	38.0	468	4 US-09-252-991A-682	Sequence 682, App
38	22.8	38.0	780	4 US-09-252-991A-729	Sequence 729, App
39	22.8	38.0	792	4 US-09-489-039A-1068	Sequence 1068, Ap
40	22.8	38.0	981	4 US-10-029-517-16	Sequence 16, Appl
41	22.8	38.0	1221	4 US-09-461-325-471	Sequence 471, App
42	22.8	38.0	1221	4 US-10-012-542-471	Sequence 471, App
43	22.8	38.0	1221	4 US-10-115-123-471	Sequence 471, App
44	22.8	38.0	1379	4 US-09-461-325-124	Sequence 124, App
45	22.8	38.0	1379	4 US-10-012-542-124	Sequence 124, App

ALIGNMENTS

```
RESULT 1
US-10-029-517-17
: Sequence 17, Application US/10029517
: Patent No. 6716627
: GENERAL INFORMATION:
: APPLICANT: Kenneth W. Dobie
: APPLICANT: Susan J. Myers
: TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
: FILE REFERENCE: RTS-0352
: CURRENT APPLICATION NUMBER: US/10/029,517
: CURRENT FILING DATE: 2001-12-20
: NUMBER OF SEQ ID NOS: 107
: SEQ ID NO 17
: LENGTH: 1804
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (73)...(1500)
: US-10-029-517-17

Query Match      45.7%; Score 27.4; DB 4; Length 1804;
Best Local Similarity 69.8%; Pred. No. 0.85;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY      7 ACCGCTCCACCTGCACACGGGTGTCACAGCGCCGACACACTGCAGCTGGCCC 59
Db      463 ACCGCCCCCAGCCAGCGGTGTCTACCTCGGCCCGACACAGCGCCGCC 515

RESULT 2
US-10-029-517-19
: Sequence 19, Application US/10029517
: Patent No. 6716627
: GENERAL INFORMATION:
: APPLICANT: Kenneth W. Dobie
: APPLICANT: Susan J. Myers
: TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
: FILE REFERENCE: RTS-0352
: CURRENT APPLICATION NUMBER: US/10/029,517
: CURRENT FILING DATE: 2001-12-20
: NUMBER OF SEQ ID NOS: 107
: SEQ ID NO 19
: LENGTH: 8186
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: unsure
: LOCATION: 6899
: OTHER INFORMATION: unknown
```

```
; NAME/KEY: unsure
; LOCATION: 7155
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7184
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7957
; OTHER INFORMATION: unknown
; NAME/KEY: intron
; LOCATION: (2997)...(3498)
; OTHER INFORMATION: intron 1
; NAME/KEY: intron:exon junction
; LOCATION: (3498)...(3499)
; OTHER INFORMATION: intron 1:exon 2
; NAME/KEY: exon
; LOCATION: (3508)...(3599)
; OTHER INFORMATION: exon 2d
; NAME/KEY: exon:intron junction
; LOCATION: (3982)...(3983)
; OTHER INFORMATION: exon 2a:intron 2a
; NAME/KEY: intron:exon junction
; LOCATION: (4205)...(4206)
; OTHER INFORMATION: intron 2c:exon 3c
; NAME/KEY: intron:exon junction
; LOCATION: (4259)...(4260)
; OTHER INFORMATION: intron 2d:exon 3d
; NAME/KEY: exon
; LOCATION: (4260)...(4328)
; OTHER INFORMATION: exon 3d
; NAME/KEY: intron:exon junction
; LOCATION: (4632)...(4633)
; OTHER INFORMATION: intron 3:exon 4
; NAME/KEY: exon
; LOCATION: (4914)...(5035)
; OTHER INFORMATION: exon 5
; NAME/KEY: intron
; LOCATION: (5266)...(6293)
; OTHER INFORMATION: intron 6
US-10-029-517-19
```

Query Match 45.7%; Score 27.4; DB 4; Length 8186;
Best Local Similarity 69.8%; Pred. No. 1.1;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
OY 7 ACCGCTCCACCTGCACACGGGGTCAACAGCGCGCCAGACACTCGACCTGCGCC 59
DB 3831 ACCGCCCCCAGCCAGCGGTGTCACTCTCGGCCCGGACACAGCGCGCGCCC 3883
```

```
RESULT 3
US-10-029-517-101
; Sequence 101, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 101
; LENGTH: 518
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-10-029-517-101
```

Query Match 43.3%; Score 26; DB 4; Length 518;
Best Local Similarity 70.0%; Pred. No. 2.2;
Matches 35; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```
OY 7 ACCGCTCCACCTGCACACGGGGTCAACAGCGCGCCAGACACTCGACCTGC 56
DB 468 ACCGCCCCCAGCCAGCGGTGTCACTCTCGGCCCGGACACAGCGCGCGC 517
```

```
RESULT 4
US-10-029-517-18
; Sequence 18, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 18
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (67)...(572)
US-10-029-517-18
```

Query Match 43.0%; Score 25.8; DB 4; Length 572;
Best Local Similarity 67.9%; Pred. No. 2.6;
Matches 36; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```
OY 7 ACCGCTCCACCTGCACACGGGGTCAACAGCGCGCCAGACACTCGACCTGCGCC 59
DB 484 ACCGCCCCCAAGCCAGCGGTGTCACTCTCGGCCCGGACACAGCGCGCGCCC 536
```

```
RESULT 5
US-10-029-517-3
; Sequence 3, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 3
; LENGTH: 1721
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (58)...(1605)
US-10-029-517-3
```

Query Match 43.0%; Score 25.8; DB 4; Length 1721;
Best Local Similarity 67.9%; Pred. No. 3.1;
Matches 36; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```
OY 7 ACCGCTCCACCTGCACACGGGGTCAACAGCGCGCCAGACACTCGACCTGCGCC 59
DB 448 ACCGCCCCCAGCCAGCGGTGTCACTCTCGGCCCGGACACAGCGCGCGCCC 500
```

```
RESULT 6
US-09-475-947A-246
; Sequence 246, Application US/09475947A
; Patent No. 6472154
; GENERAL INFORMATION:
; APPLICANT: Garner, Harold R.
; APPLICANT: Wren, Jonathan D.
; APPLICANT: Minna, John D.
```

```
; TITLE OF INVENTION: Polymorphic Repeats in Human Genes
; FILE REFERENCE: UTS0667
; CURRENT APPLICATION NUMBER: US/09/475, 947A
; CURRENT FILING DATE: 1999-12-31
; NUMBER OF SEQ ID NOS: 346
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 246
; LENGTH: 60
; TYPE: DNA
; ORGANISM: human
US-09-475-947A-246

Query Match          41.0%; Score 24.6; DB 4; Length 60;
Best Local Similarity 70.2%; Pred. No. 4.7;
Matches 33; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 7 ACCGCTCCACCTGCACACGGGGTCAACAAGCGCGCCAGACACTCGACC 53
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 13 ACCGCCCCCAGCCAGCCGCGTGTCACTCGGCCCGGACACAGGCC 59

RESULT 7
US-09-455-960-5/c
; Sequence 5, Application US/09455960
; Patent No. 6361776
; GENERAL INFORMATION:
; APPLICANT: Delcayre, Alain
; TITLE OF INVENTION: Compositions isolated from M. vaccae and
; TITLE OF INVENTION: their use in modulation of immune responses.
; FILE REFERENCE: 11000.1047
; CURRENT APPLICATION NUMBER: US/09/455, 960
; CURRENT FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 846
; TYPE: DNA
; ORGANISM: Mycobacterium vaccae
US-09-455-960-5

Query Match          40.7%; Score 24.4; DB 3; Length 846;
Best Local Similarity 63.8%; Pred. No. 8.7;
Matches 37; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGGGGTCAACAAGCGCGCCAGACACTCGACCTGGCC 58
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 487 GGACGAACCGCACCCCGCGGACCCCGGTCAACCGCGCGGAAGATGTGACCAAGCCG 430

RESULT 8
US-10-051-325-5/c
; Sequence 5, Application US/10051325
; Patent No. 6716430
; GENERAL INFORMATION:
; APPLICANT: Delcayre, Alain
; TITLE OF INVENTION: Compositions isolated from M. vaccae and
; TITLE OF INVENTION: their use in modulation of immune responses.
; FILE REFERENCE: 11000.1047c1
; CURRENT APPLICATION NUMBER: US/10/051, 325
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 09/455, 960
; PRIOR FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 846
; TYPE: DNA
; ORGANISM: Mycobacterium vaccae
US-10-051-325-5

Query Match          40.7%; Score 24.4; DB 4; Length 846;
Best Local Similarity 63.8%; Pred. No. 8.7;
Matches 37; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
```

```
QY 1 GGAAGTACCGCTCCACCTGCACACGGGGTCAACAAGCGCGCCAGACACTCGACCTGGCC 58
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 487 GGACGAACCGCACCCCGCGGACCCCGGTCAACCGCGCGGAAGATGTGACCAAGCCG 430

RESULT 9
US-09-455-960-2/c
; Sequence 2, Application US/09455960
; Patent No. 6361776
; GENERAL INFORMATION:
; APPLICANT: Delcayre, Alain
; TITLE OF INVENTION: Compositions isolated from M. vaccae and
; TITLE OF INVENTION: their use in modulation of immune responses.
; FILE REFERENCE: 11000.1047
; CURRENT APPLICATION NUMBER: US/09/455, 960
; CURRENT FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Mycobacterium vaccae
US-09-455-960-2

Query Match          40.7%; Score 24.4; DB 3; Length 2500;
Best Local Similarity 63.8%; Pred. No. 10;
Matches 37; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGGGGTCAACAAGCGCGCCAGACACTCGACCTGGCC 58
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2141 GGACGAACCGCACCCCGCGGACCCCGGTCAACCGCGCGGAAGATGTGACCAAGCCG 2084

RESULT 10
US-10-051-325-2/c
; Sequence 2, Application US/10051325
; Patent No. 6716430
; GENERAL INFORMATION:
; APPLICANT: Delcayre, Alain
; TITLE OF INVENTION: Compositions isolated from M. vaccae and
; TITLE OF INVENTION: their use in modulation of immune responses.
; FILE REFERENCE: 11000.1047c1
; CURRENT APPLICATION NUMBER: US/10/051, 325
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 09/455, 960
; PRIOR FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Mycobacterium vaccae
US-10-051-325-2

Query Match          40.7%; Score 24.4; DB 4; Length 2500;
Best Local Similarity 63.8%; Pred. No. 10;
Matches 37; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGGGGTCAACAAGCGCGCCAGACACTCGACCTGGCC 58
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2141 GGACGAACCGCACCCCGCGGACCCCGGTCAACCGCGCGGAAGATGTGACCAAGCCG 2084

RESULT 11
US-10-029-517-102
; Sequence 102, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
```


; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 102
; LENGTH: 3343
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-10-029-517-102

Query Match 40.7%; Score 24.4; DB 4; Length 3343;
Best Local Similarity 68.0%; Pred. No. 11;
Matches 34; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACCGCTCCACCTGCACACGCGGCTCACAAGCCGCCGACACTGACCTGC 56
DB 1734 ACCGCCCCCCCCAGCCCATGCTGCACCTCGCGCCGCGACACAGGCCCGC 1783

RESULT 12

US-09-489-039A-5663
; Sequence 5663, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709,2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 5663
; LENGTH: 1392
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5663

Query Match 39.3%; Score 23.6; DB 4; Length 1392;
Best Local Similarity 69.6%; Pred. No. 18;
Matches 32; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 15 ACCTGCACACGCGGCTCACAAGCGCGGACAGACTGACCTGCGCCA 60
DB 980 ACCCGCGGCGCTGTCTGCAGCGCGCGCGCTGTGTGTGCGA 1025

RESULT 13

US-09-276-531-86/c
; Sequence 86, Application US/09276531
; Patent No. 6183968
; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Reddy, Roopa
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
; TITLE OF INVENTION: RECEPTORS AND PROTEINS ASSOCIATED WITH CELL PROLIFERATION
; NUMBER OF SEQUENCES: 134
; CORRESPONDENCE ADDRESS:
ADDRESS: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/276,531
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/079,677
; FILING DATE: March 27, 1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lynn E. Murry, Ph.D.
; REGISTRATION NUMBER: 42,918
; REFERENCE/DOCKET NUMBER: PA-0008 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 86:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1772 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSTUT09
; CLONE: 1646005
US-09-276-531-86

Query Match 39.3%; Score 23.6; DB 3; Length 1772;
Best Local Similarity 64.8%; Pred. No. 19;
Matches 35; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 7 ACCGCTCCACCTGCACACGCGGCTCACAAGCGCGGACAGACTGACCTGCGCCA 60
DB 627 ACCCCTCACAGTCCACACAGGGCTTCGACCCGCGCTGCACTTCACCTTGACGCA 574

RESULT 14

US-08-956-171E-3895/c
; Sequence 3895, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:

; APPLICANT: Charles Kunsch
; APPLICANT: Gil H. Choi
; APPLICANT: Patrick S. Dillon
; APPLICANT: Craig A. Rosen
; APPLICANT: Steven C. Barash
; APPLICANT: Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman

Search completed: January 16, 2005, 03:01:43
Job time : 36.2 secs

REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 3895:
SEQUENCE CHARACTERISTICS:
LENGTH: 507 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 3895:
US-08-956-171E-3895

Query Match 38.7%; Score 23.2; DB 4; Length 507;
Best Local Similarity 61.7%; Pred. No. 21;
Matches 37; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

OY 1 GGAAGTACCGCTCCACCTGCACACGCGGCTCACAAGCGCGCCAGACACTCGACCTGCGCCA 60
Db 279 GAAACACCAACACACCGACACCGACGAGAGTGCAGAGTGCAGAACTCCAAACACCGCCA 220

RESULT 15

US-08-781-986A-3895/c
Sequence 3895, Application US/08781986A
Patent No. 6737248

GENERAL INFORMATION:

APPLICANT: Charles Kunesh
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/781,986A

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Benson, Bob

REGISTRATION NUMBER: 30,446

REFERENCE/DOCKET NUMBER: PB248PP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 3895:

SEQUENCE CHARACTERISTICS:

LENGTH: 507 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-781-986A-3895

Query Match 38.7%; Score 23.2; DB 4; Length 507;
Best Local Similarity 61.7%; Pred. No. 21;
Matches 37; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

OY 1 GGAAGTACCGCTCCACCTGCACACGCGGCTCACAAGCGCGCCAGACACTCGACCTGCGCCA 60
Db 279 GAAACACCAACACACCGACACCGACGAGAGTGCAGAGTGCAGAACTCCAAACACCGCCA 220

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 16:00:11 : Search time 170.5 Seconds
(without alignments)
1847.304 Million cell updates/sec

Title: US-10-057-136-9

Perfect score: 60

Sequence: 1 ggaagtcacgcgtccacctgc.....cagacactcgacctgcgcca 60

Scoring table:

IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	60	100.0	60	2	AAV48322	Aav48322 Nucleotid
2	52	86.7	60	2	AAV48321	Aav48321 Nucleotid
3	38.2	63.7	60	2	AAV48325	Aav48325 Nucleotid
4	36	60.0	1371	3	AAD00388	Aad00388 Human Muc
5	36	60.0	1737	3	AAD00394	Aad00394 Ubiqutin
6	34.4	57.3	525	3	AAD00385	Aad00385 Human Muc
7	34.4	57.3	891	3	AAD00391	Aad00391 Ubiqutin
8	33.4	55.7	60	2	AAV48317	Aav48317 Nucleotid
9	31.8	53.0	60	2	AAV48318	Aav48318 Nucleotid
10	31.4	52.3	93	8	ABV77125	Abv77125 Oligonuc1
11	31.4	52.3	156	8	ABV77128	Abv77128 Nucleotid
12	31.4	52.3	157	8	ABV77129	Abv77129 Nucleotid
13	31.4	52.3	157	8	ABV77127	Abv77127 Nucleotid
14	31.4	52.3	3621	8	ABV77131	Abv77131 Nucleotid
15	31.4	52.3	6877	8	ABV77133	Abv77133 Nucleotid
16	30.2	50.3	156	10	ADK68635	Adk68635 HSP65-MUC
17	30	50.0	60	2	AAV48319	Aav48319 Nucleotid
18	29	48.3	60	2	AAV48320	Aav48320 Nucleotid
19	29	48.3	60	2	AAV48324	Aav48324 Nucleotid
20	28.6	47.7	1800	10	ADK68629	Adk68629 HSP65-MUC
21	27.4	45.7	60	2	AAV48316	Aav48316 Nucleotid

C	22	27.4	45.7	309	1	AAN90579	Aan90579 pDF9.3 CD
	23	27.4	45.7	1194	12	AD157712	Ad157712 Human bre
	24	27.4	45.7	1378	12	AD157693	Ad157693 Human bre
C	25	27.4	45.7	1424	12	ADO23180	Ado23180 Antisense
	26	27.4	45.7	1428	6	ABL60159	Ab160159 Human MUC
	27	27.4	45.7	1428	12	ADO23125	Ado23125 Human MUC
	28	27.4	45.7	1457	12	ADF32627	Adf32627 Plasmid J
	29	27.4	45.7	1527	2	AAV48329	Aav48329 MiniMUC1
	30	27.4	45.7	1614	12	ADK70370	Adk70370 Respirato
	31	27.4	45.7	1630	12	AD157708	Ad157708 Human bre
	32	27.4	45.7	1634	12	AD157689	Ad157689 Human bre
	33	27.4	45.7	1712	12	AD157686	Ad157686 Human bre
	34	27.4	45.7	1738	12	AD157669	Ad157669 Human bre
	35	27.4	45.7	1755	12	AD157673	Ad157673 Human bre
	36	27.4	45.7	1774	12	AD157691	Ad157691 Plasmid J
	37	27.4	45.7	1774	12	AD157699	Ad157699 Human bre
	38	27.4	45.7	1799	12	ADO23124	Ado23124 Human MUC
	39	27.4	45.7	1803	12	AD157699	Ad157699 Human bre
	40	27.4	45.7	1804	6	ABL67539	Ab167539 Thyroid c
	41	27.4	45.7	1804	9	AAD56950	Aad56950 Human muc
	42	27.4	45.7	1804	10	ADD14719	Add14719 Human src
	43	27.4	45.7	1804	12	ADP13294	Adp13294 Renal cel
	44	27.4	45.7	1805	12	ADO28642	Ado28642 Human MUC
	45	27.4	45.7	1808	12	AD157706	Ad157706 Human bre

ALIGNMENTS

RESULT 1	AAV48322	standard; DNA; 60 BP.
ID	AAV48322	
XX	AAV48322;	
AC	20-NOV-1998	(first entry)
DT	Nucleotide sequence encoding MUC1 tandem repeat unit R7.	
DE	88; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;	
KW	tumour; tumour-associated antigen.	
KM	Homo sapiens.	
XX	WO9837095-A2.	
PN	27-AUG-1998.	
XX	24-FEB-1998;	98WO-US003693.
PF	24-FEB-1997;	97US-0038253P.
XX	(THER-) THERION BIOLOGICS CORP.	
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.	
PA	(DAND) DANA FARBER CANCER INST INC.	
PI	Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;	
XX	WPI; 1998-467492/40.	
DR	New recombinant pox virus for tumour therapy - comprises DNA encoding an	
PT	immunogenetic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.	
XX	Disclosure; Page 11; 42pp; English.	
XX	The MUC1 tandem repeat units AAV48317-V48325 were used to create an	
CC	immunogenetic mini-MUC1 fragment for inclusion in a recombinant pox virus	
CC	(RPV). The RPV was used in a pharmaceutical composition also containing	
CC	an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The	
CC	recombinant pox virus therefore encodes an immunogenetic MUC1 fragment that	
CC	does not undergo significant genetic deletion, thereby providing an	
CC	unexpectedly stable and immunogenetic pox virus. They can be used to	
CC	prevent or treat tumours expressing MUC1 tumour-associated antigens	

XX
SQ Sequence 60 BP, 14 A; 24 C; 16 G; 6 T; 0 U; 0 Other;

Query Match 100.0%; Score 60; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 3.3e-11;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGCGGGTCAACAAGCGGCCAGACACTCGACCTGGCGCA 60
DB 1 GGAAGTACCGCTCCACCTGCACACGCGGGTCAACAAGCGGCCAGACACTCGACCTGGCGCA 60

RESULT 2

ID AAV48321 standard; DNA; 60 BP.
XX
AC AAV48321;

DT 20-NOV-1998 (first entry)

DE Nucleotide sequence encoding MUC1 tandem repeat unit R6.

KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW tumour; tumour-associated antigen.

OS Homo sapiens.

PN WO9837095-A2.

PD 27-AUG-1998.

PF 24-FEB-1998; 98WO-US003693.

PR 24-FEB-1997; 97US-0038253P.

PA (THER-) THERION BIOLOGICS CORP.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (DAND) DANA FARBER CANCER INST INC.

PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;

DR WPI; 1998-467492/40.

PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS Disclosure; Page 11; 42pp; English.

CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens

SQ Sequence 60 BP; 13 A; 27 C; 17 G; 3 T; 0 U; 0 Other;

Query Match 86.7%; Score 52; DB 2; Length 60;
Best Local Similarity 91.7%; Pred. No. 1.9e-08;
Matches 55; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGCGGGTCAACAAGCGGCCAGACACTCGACCTGGCGCA 60
DB 1 GGCAGCACCGCACCGCCCGCACACGGGTCTACAAAGCGGCCAGACACTCGACCTGGCGCA 60

RESULT 3

ID AAV48325 standard; DNA; 60 BP.

XX
AC AAV48325;

XX
DT 20-NOV-1998 (first entry)

DE Nucleotide sequence encoding MUC1 tandem repeat unit R10.

KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW tumour; tumour-associated antigen.

OS Homo sapiens.

PN WO9837095-A2.

PD 27-AUG-1998.

PF 24-FEB-1998; 98WO-US003693.

PR 24-FEB-1997; 97US-0038253P.

PA (THER-) THERION BIOLOGICS CORP.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (DAND) DANA FARBER CANCER INST INC.

PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;

DR WPI; 1998-467492/40.

PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS Disclosure; Page 11; 42pp; English.

CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens

SQ Sequence 60 BP; 12 A; 22 C; 19 G; 7 T; 0 U; 0 Other;

Query Match 63.7%; Score 38.2; DB 2; Length 60;
Best Local Similarity 78.0%; Pred. No. 0.0011;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGCGGGTCAACAAGCGGCCAGACACTCGACCTGGCGCC 59
DB 1 GGTAAGTACAGCGCCACCGCCGACATGGGTCTACGAGCGCTCCGGATACGAGACCGCGGCC 59

RESULT 4

ID AAD00388 standard; DNA; 1371 BP.

AC AAD00388;

DT 29-AUG-2000 (first entry)

DE Human Mucin 1 (MUC-1) protein fragment encoding DNA #5.

KW Human; Mucin 1; MUC-1; tumour; pMR330 expression vector; anti-tumour;
KW therapy; immune response; cytostatic; vaccine; ds.

OS Homo sapiens.

Key Location/Qualifiers
CDS 1..1371
/*tag= a
/product= "MUC-1 protein fragment"

PN WO200025827-A2.

PD 11-MAY-2000.
XX
XX 18-OCT-1999; 99WO-EP007874.
PF
XX 30-OCT-1998; 98IT-MI002330.
PR
XX (MENA) MENARINI RICERCHE SPA.
PA
XX Parente D, Di Massimo AM, De Santis R;
PI
XX WPI; 2000-365410/31.
DR P-PSDB; AAY71024.
XX
XX Composition containing one or more DNA molecules encoding fragments of a
PT Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
PT tumor therapy.
XX
XX Claim 16; Fig 5; 56pp; English.
XX
XX The present sequence is a DNA encoding a fragment of human Mucin 1 (MUC-
CC 1) antigenic protein which is overexpressed in tumour cells. The sequence
CC was obtained by PCR from plasmids pMRS166, pMRS167, pMRS168 and pMRS169
CC which contain MUC-1 DNA from BT20 tumour cells. It corresponds to
CC nucleotides 136-1497 of the EMBL sequence J05581 with a start codon and
CC two stop codons. The present sequence is cloned into a pMRS30 expression
CC vector and used in pharmaceutical composition e.g. vaccine for inducing
CC an antigen-specific anti-tumour immune response. Composition containing
CC this DNA molecule is useful in anti-tumour therapy of patients affected
CC with tumours characterised by high MUC-1 expression
XX
XX Sequence 1371 BP; 300 A; 450 C; 321 G; 300 T; 0 U; 0 Other;
SQ
Query Match 60.0%; Score 36; DB 3; Length 1371;
Best Local Similarity 75.0%; Pred. No. 0.0092;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 GGAAGTACCGCTCCACCTGCACACGGGGTCAACAAGCGCCAGACACTGCACCTGCGCCA 60
DB 325 GGAAGTACCGCTCCACCGACACACGGTGTACCTCGGCTCCGATACCAAGCGGCCCA 384
RESULT 5
AAD00394
ID AAD00394 standard; DNA; 1737 BP.
XX
AC AAD00394;
XX
DT 15-SEP-2003 (revised)
DT 29-AUG-2000 (first entry)
XX
DE Ubiquitin-E. coli LacI-human Mucin 1 fusion protein encoding DNA #5.
XX
KM Ubiquitin; LacI; beta-galactosidase; fusion protein; human; Mucin 1;
KM MUC-1; tumour; pMRS30 expression vector; anti-tumour; therapy;
KM immune response; cytosstatic; vaccine; ds.
XX
OS Homo sapiens.
OS Escherichia coli.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 1..1737
FT /*tag= a
FT /product= "Ubiquitin-LacI-MUC-1 fusion protein"
FT 1..369
FT /*tag= b
FT /label= UBILacI_DNA
FT /note= "Includes ubiquitin-E. coli LacI fusion DNA"
FT 370..1737
FT /*tag= c
FT /note= "Human MUC-1 partial DNA that corresponds to
FT nucleotides 136-1497 of the EMBL sequence J05581 with two
FT stop codons"

XX
PN WO200025827-A2.
XX
XX 11-MAY-2000.
PD
XX 18-OCT-1999; 99WO-EP007874.
PF
XX 30-OCT-1998; 98IT-MI002330.
PR
XX (MENA) MENARINI RICERCHE SPA.
PA
XX Parente D, Di Massimo AM, De Santis R;
PI
XX WPI; 2000-365410/31.
DR P-PSDB; AAY71030.
XX
XX Composition containing one or more DNA molecules encoding fragments of a
PT Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
PT tumor therapy.
XX
XX Claim 18; Fig 11; 56pp; English.
XX
XX The present sequence is a DNA encoding a fusion protein consisting of
CC human Mucin 1 (MUC-1) fragment fused to UBILacI sequence at the N-
CC terminus. The UBILacI sequence consists of ubiquitin from MCF7 cell line
CC and a portion of E. coli beta-galactosidase (lacI). MUC-1 is an antigenic
CC protein overexpressed in tumour cells. The present sequence is cloned
CC into a pMRS30 expression vector and used in pharmaceutical composition
CC e.g. vaccine for inducing an antigen-specific anti-tumour immune
CC response. Composition containing this DNA molecule is useful in anti-
CC tumour therapy of patients affected with tumours characterised by high
CC MUC-1 expression. (Updated on 15-SEP-2003 to standardise OS field)
XX
XX Sequence 1737 BP; 392 A; 558 C; 418 G; 369 T; 0 U; 0 Other;
SQ
Query Match 60.0%; Score 36; DB 3; Length 1737;
Best Local Similarity 75.0%; Pred. No. 0.0094;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 GGAAGTACCGCTCCACCTGCACACGGGGTCAACAAGCGCCAGACACTGCACCTGCGCCA 60
DB 691 GGAAGTACCGCTCCACCGACACACGGTGTACCTCGGCTCCGATACCAAGCGGCCCA 750
RESULT 6
AAD00385
ID AAD00385 standard; DNA; 525 BP.
XX
AC AAD00385;
XX
DT 29-AUG-2000 (first entry)
DT
XX
DE Human Mucin 1 (MUC-1) protein fragment encoding DNA #2.
XX
KM Human; Mucin 1; MUC-1; tumour; pMRS30 expression vector; anti-tumour;
KM therapy; immune response; cytosstatic; vaccine; ds.
XX
OS Homo sapiens.
OS
OS
XX
FH Key Location/Qualifiers
FT CDS 1..525
FT /*tag= a
FT /product= "MUC-1 protein fragment".
FT
PN WO200025827-A2.
XX
XX 11-MAY-2000.
PD
XX 18-OCT-1999; 99WO-EP007874.
PF
XX 30-OCT-1998; 98IT-MI002330.
PR
XX (MENA) MENARINI RICERCHE SPA.
PA

```
XX
PI Parente D, Di Massimo AM, De Santis R;
XX
XX WPI; 2000-365410/31.
DR
DR P-PSDB; AAY71021.
XX
XX Composition containing one or more DNA molecules encoding fragments of a
PT Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
PT tumor therapy.
XX
XX Claim 16; Fig 2; 56pp; English.
XX
XX The present sequence is a DNA encoding a fragment of human Mucin 1 (MUC-
CC 1) antigenic protein which is overexpressed in tumour cells. The sequence
CC was obtained from BT20 tumour cells by reverse transcriptase- PCR and
CC corresponds to nucleotides 205-720 of the EMBL sequence J05581 with a
CC start codon and two stop codons. The present sequence is cloned into a
CC pMRS30 expression vector and used in pharmaceutical composition e.g.
CC vaccine for inducing an antigen-specific anti-tumour immune response.
CC Composition containing this DNA molecule is useful in anti-tumour therapy
CC of patients affected with tumours characterised by high MUC-1 expression
XX
SQ Sequence 525 BP; 111 A; 199 C; 128 G; 87 T; 0 U; 0 Other;
Query Match 57.3%; Score 34.4; DB 3; Length 525;
Best Local Similarity 73.3%; Pred. No. 0.029;
Matches 44; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
OY 1 GGAAGTACCGCTCCACCTGCACACGCGGTCAACAAGCGCGCCAGACACTCGACCTGGCCCA 60
Db 256 GGAAGTACTGCTCCACACGACACACGCGTGTACTCCTCGGCTCCGGATACCAAGCGCGCCCA 315
RESULT 7
AAD00391
ID AAD00391 standard; DNA; 891 BP.
XX
XX AAD00391;
AC
XX 15-SEP-2003 (revised)
DT 29-AUG-2000 (first entry)
XX
DE Ubiquitin-E. coli LacI-human Mucin 1 fusion protein encoding DNA #2.
XX
KW Ubiquitin; LacI; beta-galactosidase; fusion protein; human; Mucin 1;
KW MUC-1; tumour; pMRS30 expression vector; anti-tumour; therapy;
KW immune response; cytostatic; vaccine; ds.
XX
OS Homo sapiens.
OS Escherichia coli.
OS Chimeric.
XX
XX Key Location/Qualifiers
FH CDS 1..891
FT /*tag= a
FT /product= "Ubiquitin-LacI-MUC-1 fusion protein"
FT 1..369
FT /*tag= b
FT /label= UBILacI DNA
FT /note= "Includes ubiquitin-E. coli LacI fusion DNA"
FT 370..891
FT /*tag= c
FT /note= "Human MUC-1 partial DNA that corresponds to
FT nucleotides 205-720 of the EMBL sequence J05581 with two
FT stop codons"
XX
XX WO200025827-A2.
XX
XX 11-MAY-2000.
XX
XX 18-OCT-1999; 99WO-EP007874.
XX
XX 30-OCT-1998; 98IT-MI002330.
```

```
XX
XX (MENA ) MENARINI RICERCHE SPA.
XX
XX Parente D, Di Massimo AM, De Santis R;
PI
XX
XX WPI; 2000-365410/31.
DR
DR P-PSDB; AAY71027.
XX
XX Composition containing one or more DNA molecules encoding fragments of a
PT Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
PT tumor therapy.
XX
XX Claim 18; Fig 8; 56pp; English.
XX
XX The present sequence is a DNA encoding a fusion protein consisting of
CC human Mucin 1 (MUC-1) fragment fused to UBILacI sequence at the N-
CC terminus. The UBILacI sequence consists of ubiquitin from MCF7 cell line
CC and a portion of E. coli beta-galactosidase (LacI). MUC-1 is an antigenic
CC protein overexpressed in tumour cells. The present sequence is cloned
CC into a pMRS30 expression vector and used in pharmaceutical composition
CC e.g. vaccine for inducing an antigen-specific anti-tumour immune
CC response. Composition containing this DNA molecule is useful in anti-
CC tumour therapy of patients affected with tumours characterised by high
CC MUC-1 expression. (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ Sequence 891 BP; 203 A; 307 C; 225 G; 156 T; 0 U; 0 Other;
Query Match 57.3%; Score 34.4; DB 3; Length 891;
Best Local Similarity 73.3%; Pred. No. 0.031;
Matches 44; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
OY 1 GGAAGTACCGCTCCACCTGCACACGCGGTCAACAAGCGCGCCAGACACTCGACCTGGCCCA 60
Db 622 GGAAGTACTGCTCCACACGACACACGCGTGTACTCCTCGGCTCCGGATACCAAGCGCGCCCA 681
RESULT 8
AAV48317
ID AAV48317 standard; DNA; 60 BP.
XX
XX AAV48317;
AC
XX 20-NOV-1998 (first entry)
DT
XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit R2.
XX
KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW tumour; tumour-associated antigen.
XX
OS Homo sapiens.
XX
XX WO9837095-A2.
XX
XX 27-AUG-1998.
PD
XX 24-FEB-1998; 98WO-US003693.
PF
XX 24-FEB-1997; 97US-0038253P.
PR
XX 24-FEB-1997; 97US-0038253P.
XX
XX (THER-) THERION BIOLOGICS CORP.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX (DAND ) DANA FARBER CANCER INST INC.
XX
XX Schlom J, Kantov J, Kufe D, Panicali D, Gritz L;
XX
XX WPI; 1998-467492/40.
XX
XX New recombinant pox virus for tumour therapy - comprises DNA encoding an
XX immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX Disclosure; Page 11; 42pp; English.
XX
XX The MUC1 tandem repeat units AAV48317-V48325 were used to create an
```


CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ Sequence 60 BP; 17 A; 21 C; 13 G; 9 T; 0 U; 0 Other;

Query Match 55.7%; Score 33.4; DB 2; Length 60;
Best Local Similarity 72.9%; Pred. No. 0.05;
Matches 43; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGGGGTCAACAAGCGGCCAGACACTGCACTGCGGCC 59
Db 1 GGAGTACTGCACCGCACCGGCATGCGTAACTACAGCACTGATACAGAAGCCTGCACC 59

RESULT 9
AAV48318
ID AAV48318 standard; DNA; 60 BP.
XX
AC AAV48318;
XX
DT 20-NOV-1998 (first entry)
XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit R3.
XX
KM 88; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KM tumour; tumour-associated antigen.
XX
OS Homo sapiens.
XX
PN WO9837095-A2.
XX
PD 27-AUG-1998.
XX
PF 24-FEB-1998; 98WO-US003693.
XX
PR 24-FEB-1997; 97US-0038253P.
XX
PA (THER-) THERION BIOLOGICS CORP.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (DAND) DANA FARBER CANCER INST INC.
XX
PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX
DR WPI; 1998-467492/40.
XX
PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS Disclosure; Page 11; 42pp; English.
XX
CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ Sequence 60 BP; 7 A; 29 C; 19 G; 5 T; 0 U; 0 Other;

Query Match 53.0%; Score 31.8; DB 2; Length 60;
Best Local Similarity 71.2%; Pred. No. 0.18;
Matches 42; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGGGGTCAACAAGCGGCCAGACACTGCACTGCGGCC 59
Db 1 GGATCCACCGCGCGCGCTGCGCAGGAGTGAAGTCGCGGCCCGGACACGCGCCCGCTCC 59

RESULT 10
ABV77125
ID ABV77125 standard; DNA; 93 BP.
XX
AC ABV77125;
XX
DT 28-MAR-2003 (first entry)
XX
DE Oligonucleotide used to create a plasmid insert encoding MUC1 epitope.
XX
KM Intracellular analysis; intracellular interaction; intrabody; ScFv; MUC1;
KM mucin 1; 88.
XX
OS Synthetic.
XX
PN WO200286505-A2.
XX
PD 31-OCT-2002.
XX
PF 02-APR-2002; 2002WO-GB001235.
XX
PR 31-MAR-2001; 2001GB-00008165.
XX
PA (UYMA-) UNIV VICTORIA MANCHESTER.
XX
PI Roderick SPB;
XX
DR WPI; 2003-103417/09.
XX
PT Intracellularly analyzing or detecting the presence of a target molecule
PT within a biological cell, by determining the presence or signal
PT representative of binding of polypeptide species to the target molecule.
XX
PS Example; Page 19; 51pp; English.
XX
CC The specification describes a method for the intracellular analysis of a
CC target molecule within a biological cell. The method comprises expressing
CC within the cell a first polypeptide sequence comprised of a first binding
CC species capable of binding to the target molecule and a first reporter
CC moiety attached to the first binding species; expressing within the cell
CC a second polypeptide sequence comprised of a second binding species
CC capable of competing with the target molecule for binding species
CC of competing with the target molecule for binding of the first binding
CC species and a second reporter moiety, the first and second reporter
CC moieties being such that on binding together of the first and second
CC binding species the first and second reporter moieties interact so as to
CC be capable of producing a signal that can be differentiated from one
CC not interact; and effecting a measurement to determine the presence or
CC signal representative of binding of the first and second binding species.
CC The method is useful for the intracellular analysis of a target molecule,
CC e.g. for detecting the presence and/or amount of target molecules in
CC cells. The method is also useful for studying the interactions of
CC intracellular molecules. Oligonucleotides ABV77125-26 were used to create
CC a plasmid insert encoding the MUC1 epitope of human mucin 1. The
CC oligonucleotides were used in an intracellular assay for MUC1,
CC demonstrating the method of the invention. In this assay, the first
CC polypeptide is an anti-MUC1 ScFv, and the second polypeptide comprises
CC the MUC1 epitope attached to yellow fluorescent protein
XX
SQ Sequence 93 BP; 20 A; 41 C; 19 G; 13 T; 0 U; 0 Other;

Query Match 52.3%; Score 31.4; DB 8; Length 93;
Best Local Similarity 85.4%; Pred. No. 0.26;
Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGGGGTCAACAAGCGCGGCC 41
Db 38 GGATTAACCGCTCTCTGCGCCAGGAGTCAACAAGCGCACCC 78

```
RESULT 11
ABV77128
ID ABV77128 standard; DNA, 156 BP.
XX
XX
AC ABV77128;
XX
XX
DT 28-MAR-2003 (first entry)
XX
DE Nucleotide sequence of a mutated plasmid insert encoding MUC1 epitope.
XX
XX
KM Intracellular analysis; intracellular interaction; intrabody; ScFv; MUC1;
XX mucin 1; ss.
XX
OS Synthetic.
XX
FH Key location/Qualifiers
FT CDS 11..115
FT /*tag= a
FT /note= "partial sequence"
XX
XX
PN WO200286505-A2.
XX
XX
PD 31-OCT-2002.
XX
XX
PF 02-APR-2002; 2002WO-GB001235.
XX
XX
PR 31-MAR-2001; 2001GB-00008165.
XX
XX
PA (UYMA-) UNIV VICTORIA MANCHESTER.
XX
PI Roderick SPB;
XX
XX
DR WPI; 2003-103417/09.
DR P-PSDB; ABB99633.
XX
XX
PT Intracellularly analyzing or detecting the presence of a target molecule
PT within a biological cell, by determining the presence or signal
PT representative of binding of polypeptide species to the target molecule.
XX
XX
PS Example; Page 20-21; 51pp; English.
XX
XX
CC The specification describes a method for the intracellular analysis of a
CC target molecule within a biological cell. The method comprises expressing
CC within the cell a first polypeptide sequence comprised of a first binding
CC species capable of binding to the target molecule and a first reporter
CC moiety attached to the first binding species; expressing within the cell
CC a second polypeptide sequence comprised of a second binding species
CC capable of competing with the target molecule for binding within the cell
CC of competing with the target molecule for binding of the first binding
CC species and a second reporter moiety, the first and second reporter
CC moieties being such that on binding together of the first and second
CC binding species the first and second reporter moieties interact so as to
CC be capable of producing a signal that can be differentiated from one
CC capable of being generated when the first and second reporter moieties do
CC not interact; and effecting a measurement to determine the presence or
CC signal representative of binding of the first and second binding species.
CC The method is useful for the intracellular analysis of a target molecule,
CC e.g. for detecting the presence and/or amount of target molecules in
CC cells. The method is also useful for studying the interactions of
CC intracellular molecules. The present sequence encodes a mutated MUC1
CC epitope of human mucin 1 and its surrounding amino acids in a tandem
CC repeat. It was used in an intracellular assay for MUC1, demonstrating the
CC method of the invention. In this assay, the first polypeptide is an anti-
CC MUC1 ScFv, and the second polypeptide comprises the MUC1 epitope attached
CC to yellow fluorescent protein
XX
SQ Sequence 156 BP; 32 A; 69 C; 33 G; 22 T; 0 U; 0 Other;
```

Query Match 52.3%; Score 31.4; DB 8; Length 156;
Best Local Similarity 85.4%; Pred. No. 0.27;
Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

1 GGAAGTACCGCTCCACCTGCACACGGGGTCAACAGCGGCC 41

```
DB 38 GGAATCACCGCTCCTCCTGCGCCAGAGTCACACGCGACC 78
|||||
RESULT 12
ABV77129
ID ABV77129 standard; DNA, 157 BP.
XX
XX
AC ABV77129;
XX
XX
DT 28-MAR-2003 (first entry)
XX
XX
DE Nucleotide sequence of a plasmid insert encoding MUC1 epitope.
XX
XX
KM Intracellular analysis; intracellular interaction; intrabody; ScFv; MUC1;
XX mucin 1; ss.
XX
OS Synthetic.
XX
FH Key location/Qualifiers
FT CDS 11..157
FT /*tag= a
FT /note= "partial sequence"
XX
XX
PN WO200286505-A2.
XX
XX
PD 31-OCT-2002.
XX
XX
PF 02-APR-2002; 2002WO-GB001235.
XX
XX
PR 31-MAR-2001; 2001GB-00008165.
XX
XX
PA (UYMA-) UNIV VICTORIA MANCHESTER.
XX
PI Roderick SPB;
XX
XX
DR WPI; 2003-103417/09.
DR P-PSDB; ABB99634.
XX
XX
PT Intracellularly analyzing or detecting the presence of a target molecule
PT within a biological cell, by determining the presence or signal
PT representative of binding of polypeptide species to the target molecule.
XX
XX
PS Example; Page 21-22; 51pp; English.
XX
XX
CC The specification describes a method for the intracellular analysis of a
CC target molecule within a biological cell. The method comprises expressing
CC within the cell a first polypeptide sequence comprised of a first binding
CC species capable of binding to the target molecule and a first reporter
CC moiety attached to the first binding species; expressing within the cell
CC a second polypeptide sequence comprised of a second binding species
CC capable of competing with the target molecule for binding within the cell
CC of competing with the target molecule for binding of the first binding
CC species and a second reporter moiety, the first and second reporter
CC moieties being such that on binding together of the first and second
CC binding species the first and second reporter moieties interact so as to
CC be capable of producing a signal that can be differentiated from one
CC capable of being generated when the first and second reporter moieties do
CC not interact; and effecting a measurement to determine the presence or
CC signal representative of binding of the first and second binding species.
CC The method is useful for the intracellular analysis of a target molecule,
CC e.g. for detecting the presence and/or amount of target molecules in
CC cells. The method is also useful for studying the interactions of
CC intracellular molecules. The present sequence encodes the MUC1 epitope of
CC human mucin 1 and its surrounding amino acids in a tandem repeat. It was
CC used in an intracellular assay for MUC1, demonstrating the method of the
CC invention. In this assay, the first polypeptide is an anti-MUC1 ScFv, and
CC the second polypeptide comprises the MUC1 epitope attached to yellow
XX
SQ Sequence 157 BP; 32 A; 68 C; 35 G; 22 T; 0 U; 0 Other;
```

Query Match 52.3%; Score 31.4; DB 8; Length 157;

Query Match 52.3%; Score 31.4; DB 8; Length 3621;
Best Local Similarity 85.4%; Pred. No. 0.4;
Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 GGAAGTACCGCTCCACCTGCACACGGGGTCAAGCGCGCC 41
DB 659 GGATCTACCGCTCTCTCTGCCCCAGAGTCACAAGCGCAC 699

QY 1 GGAAGTACCGCTCCACCTGCACACGGGGTCAAGCGCGCC 41
DB 701 GGATCTACCGCTCTCTCTGCCCCAGAGTCACAAGCGCAC 741
Search completed: January 15, 2005, 20:36:11
Job time : 171.5 secs

RESULT 15

ABV77133
ID ABV77133 standard; DNA; 6877 BP.
XX
AC ABV77133;
XX
DT 28-MAR-2003 (first entry)
XX
DE Nucleotide sequence of plasmid pBudMUC-EYFPscFVECFP.
XX
KW Intracellular analysis; intracellular interaction; intrabody; ScFv; MUC1;
XX mucin 1; pBudMUC-EYFPscFVECFP; ss.
OS Synthetic.
XX
XX WO200286505-A2.
XX
PD 31-OCT-2002.
XX
XX
PF 02-APR-2002; 2002WO-GB001235.
XX
XX
PR 31-MAR-2001; 2001GB-00008165.
XX
XX
PA (UWMA-) UNIV VICTORIA MANCHESTER.
XX
PI Roderick SPB;
XX
DR WPI; 2003-103417/09.
XX
XX
PT Intracellularly analyzing or detecting the presence of a target molecule
PT within a biological cell, by determining the presence or signal
PT representative of binding of polypeptide species to the target molecule.
XX
XX
PS Example; Page 37-40; 51pp; English.
XX
XX

CC The specification describes a method for the intracellular analysis of a
CC target molecule within a biological cell. The method comprises expressing
CC within the cell a first polypeptide sequence comprised of a first binding
CC species capable of binding to the target molecule and a first reporter
CC moiety attached to the first binding species; expressing within the cell
CC a second polypeptide sequence comprised of a second binding species
CC capable of competing with the target molecule for binding species capable
CC of competing with the target molecule for binding of the first binding
CC species and a second reporter moiety, the first and second reporter
CC moieties being such that on binding together of the first and second
CC binding species the first and second reporter moieties interact so as to
CC be capable of producing a signal that can be differentiated from one
CC not interact; and effecting a measurement to determine the presence or
CC signal representative of binding of the first and second binding species.
CC The method is useful for the intracellular analysis of a target molecule,
CC e.g. for detecting the presence and/or amount of target molecules in
CC intracellular molecules. The present sequence represents a plasmid used
CC in an intracellular assay for the MUC1 epitope of human mucin1,
CC demonstrating the method of the invention. In this assay, the first
CC polypeptide is an anti-MUC1 ScFv, and the second polypeptide comprises
CC the MUC1 epitope attached to yellow fluorescent protein
XX

SQ Sequence 6877 BP; 1511 A; 1962 C; 1964 G; 1440 T; 0 U; 0 Other;

Query Match 52.3%; Score 31.4; DB 8; Length 6877;
Best Local Similarity 85.4%; Pred. No. 0.43;
Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 17:40:52 : Search time 722.8 Seconds
(without alignments)
3925.542 Million cell updates/sec

Title: US-10-057-136-9
Perfect score: 60
Sequence: 1 ggaagtcacgcctccacactgc.....cagacactcgactcgcgcca 60

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	2297	6	BD272907 A recombi
2	36	60.0	1371	6	BD225144 Medicinal
3	36	60.0	1737	6	BD225150 Medicinal
4	34.4	57.3	525	6	BD225141 Medicinal
5	34.4	57.3	891	6	BD225147 Medicinal
6	32.8	54.7	132	6	AX180275 Sequence
7	32.8	54.7	144	6	AX180275 Sequence
8	31.4	52.3	93	6	AX703425 Sequence
9	31.4	52.3	156	6	AX703427 Sequence
10	31.4	52.3	157	6	AX703428 Sequence
11	31.4	52.3	157	6	AX703431 Sequence
12	31.4	52.3	3621	6	AX703433 Sequence
13	31.4	52.3	6877	6	AX703435 Sequence
14	31.2	52.0	132	6	AX180282 Sequence
15	31.2	52.0	144	6	AX180281 Sequence
16	30	50.0	736	10	AF484219 Cavia por
17	29	48.3	303400	1	AP006578 Gloeobact
18	28	46.7	32748	1	AB070951 Streptomy
19	28	46.7	296300	1	AP005035 Streptomy

C	20	27.4	45.7	120	6	AX192396	AX192396 Sequence
C	21	27.4	45.7	120	6	BD000571	BD000571 Human pol
	22	27.4	45.7	1414	12	AF423031	AF423031 Synthetic
	23	27.4	45.7	1455	6	CQ715242	CQ715242 Sequence
	24	27.4	45.7	1457	6	AX959914	AX959914 Sequence
	25	27.4	45.7	1774	6	AX959684	AX959684 Sequence
	26	27.4	45.7	1774	6	AX959912	AX959912 Sequence
	27	27.4	45.7	1804	6	AR492318	AR492318 Sequence
	28	27.4	45.7	1804	6	AX335367	AX335367 Sequence
	29	27.4	45.7	1804	9	HUMMUCAB	J05581 Human polym
	30	27.4	45.7	1834	12	AF423030	AF423030 Synthetic
	31	27.4	45.7	1835	6	AX959918	AX959918 Sequence
	32	27.4	45.7	2135	6	AX959916	AX959916 Sequence
	33	27.4	45.7	4139	6	CQ834017	CQ834017 Sequence
	34	27.4	45.7	4139	6	AX334899	AX334899 Sequence
	35	27.4	45.7	4139	6	AX335372	AX335372 Sequence
	36	27.4	45.7	4139	6	AX336712	AX336712 Sequence
	37	27.4	45.7	4139	6	AX409474	AX409474 Sequence
	38	27.4	45.7	4139	6	AX440481	AX440481 Sequence
	39	27.4	45.7	4139	9	HUMPANNU	J05582 Human pancr
	40	27.4	45.7	7188	9	AY463543	AY463543 Homo sapi
	41	27.4	45.7	8181	6	AX406624	AX406624 Sequence
	42	27.4	45.7	8181	9	HUMPEM	M61170 Human polym
	43	27.4	45.7	8186	6	AR492320	AR492320 Sequence
C	44	27.4	45.7	133525	9	AL713999	AL713999 Human DNA
C	45	26.8	44.7	216050	1	AL646076	AL646076 Ralstonia

ALIGNMENTS

RESULT 1	BD272907	2297 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD272907	A recombinant vector expressing multiple constitutulatory molecules and uses thereof.			
DEFINITION	BD272907	A recombinant vector expressing multiple constitutulatory molecules and uses thereof.			
ACCESSION	BD272907.1	GI:33082675			
VERSION	BD272907				
KEYWORDS	JP 2002531133-A/1.				
SOURCE	synthetic construct				
ORGANISM	synthetic construct				
REFERENCE	artificial sequences.				
AUTHORS	1 (bases 1 to 2297)				
TITLE	Schlom,J., Hodge,J. and Panicali,D.				
JOURNAL	A recombinant vector expressing multiple constitutulatory molecules and uses thereof				
COMMENT	Patent: JP 2002531133-A 1 24-SEP-2002; THE UNITED STATES OF AMERICA, THERION BIOLOGICS CORP				
	OS Artificial Sequence				
	PN JP 2002531133-A/1				
	PD 24-SEP-2002				
	PF 12-NOV-1999 JP 2000586927				
	PR 09-DEC-1998 US 60/111582				
	PI JEFFREY SCHLOM, JAMES HODGE, DENNIS PANICALI				
	PC C12N15/02, A61K35/12, A61K35/74, A61K35/76, A61K38/00, A61K39/00,				
	PC A61K39/12,				
	PC A61K39/125, A61K39/21, A61K39/235, A61K39/245, A61K39/275, A61K39/				
	PC 29, A61K48/00,				
	PC A61P1/04, A61P31/04, A61P31/10, A61P31/12, A61P35/00, A61P37/02, PC				
	A61P37/06,				
	PC C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12N7/00, C12Q1/02, G01N33/				
	PC 53, C12N15/00,				
	PC C12N5/00, A61K37/02				
	CC Description of Artificial Sequence: VECTOR SEQUENCE FH Key				
	FT Location/Qualifiers				
	1. .2297				
FEATURES	source				
	location/Qualifiers				
	1. .2297				
	/organism='Artificial Sequence'.				
	/organism="synthetic construct"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:32630"				
ORIGIN					

Query Match	100.0%;	Score 60;	DB 6;	Length 2297;
Best Local Similarity	100.0%;	Pred. No. 1e-09;		
Matches	60;	Conservative	0;	Mismatches 0.

QY 1 GGAAGTACCGCTCCACTGCACA CGGGGTCA CAGA GCGCGCAGACATTCGACCTGGGCCA 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 706 GGAAGTACCGCTCCACTGCACA CGGGGTCA CAGA GCGCGCAGACATTCGACCTGGGCCA 765

RESULT 2	
BD225144	
LOCUS	
DEFINITION	BD225144
	Medicinal composition having antitumor effect and containing DNA encoding antigenic protein.
ACCESSION	1371 bp DNA linear PAT 17-JUL-2003

REFERENCE
AUTHORS
TITLE
JOURNAL

```
FEATURES
source      /organism='Homo sapiens (human)'.
            Location/Qualifiers
            1..1371
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
```

Query Match	60.0%;	Score 36;	DB 6;	Length 1371;
Best Local Similarity	75.0%;	Pred. NO. 0.14;		
Matches 45;	Conservative 0;	Mismatches 15;		

[illegible]

RESULT 3	
BD225150	
LOCUS	
DEFINITION	BD225150 1737 bp DNA linear PAT 17-JUL-2003
ACCESSION	Medicinal composition having antitumor effect and containing DNA
DESCRIPTION	encoding antigenic protein.

ORGANISM	REFERENCE	AUTHORS
Homo sapiens		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
1 (bases 1 to 1737)		
Palente, D., Massimo, A.M.D. and Desantis, R.		

TITLE	Medicinal composition having antitumor effect and containing DNA encoding antigenic protein
JOURNAL	Patent: JP 2002528519-A 11 03-SEP-2002:

COMMENT	OS	Homo sapiens (human)
	PN	JP 2002528519-A/11
	PD	03-SEP-2002
	PF	18-OCT-1999 JP 2000579265
	PR	30-OCT-1998 IT MI98A002330
	PI	DINO PALLENTE, ANNA MARIA D MASSIMO, RITA DESANTIS PC
		A61K38/00, A61K35/76, A61K39/00, A61K48/00, A61P35/00, C12N15/09, PC
		A61K37/02,

```

FEATURES
source      location/Qualifiers
1. 1737     /organism='Homo sapiens (human)'.
           /organism='Homo sapiens'
           /mol_type='genomic DNA'
           /db_xref='taxon:9606'

```

Query Match	60.0%;	Score 36;	DB 6;	Length 1737;
Best Local Similarity	75.0%;	Pred. No. 0.13;		
Matches	45;	Conservative	0;	Mismatches 15.
				Indels 0.
				Cons 0.

QY 1 GGAAGTACCGGCTCCACCCTGCACACGGGGGTTCACAAGCGCGCCAGACACTCGACCTGCCCA 60
Db 691 GGAAGTACCGGCTCCACCAGCACACGGTGTTACTCTCGGCTCCGGATACCAAGCGCGGCCCA 75

RESULT 4	BD225141	LOCUS	DEFINITION
	BD225141	525 bp	DNA linear
			Medicinal composition having antitumor effect and containing DNA encoding antigenic protein.
ACCESION			PAT 17-JUL-2001

REFERENCE	AUTHORS	TITLE	JOURNAL
1 (bases 1 to 525)	Paliente, D., Massimo, A.M.D. and Desantis, R.	Medicinal composition having antitumor effect and containing DNA encoding antigenic protein	Patent: JP 2002528519-A 2 03-SHP-2002; MEMORATI BICCHIERI and

```

FEATURES
source
1.525
Location/Qualifiers
/organism="Homo sapiens (human)"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```


ORIGIN

Query Match 57.3%; Score 34.4; DB 6; Length 525;
Best Local Similarity 73.3%; Pred. No. 0.52;
Matches 44; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGGGGTCAACAAGCCGCCAGACACTGCACCTGCGCCA 60
|||||
Db 256 GGAAGTACTGCTCCACCAACACACACGGTGTACCTCGGCTCCGATACCAAGCCGCCCA 315

RESULT 5

BD225147 891 bp DNA linear PAT 17-JUL-2003

LOCUS BD225147 Medicinal composition having antitumor effect and containing DNA
encoding antigenic protein.

ACCESSION BD225147

VERSION BD225147.1 GI:33034917

KEYWORDS JP 2002528519-A/8.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.

1 (bases 1 to 891)

Palente,D., Massimo,A.M.D. and Desantis,R.

Medicinal composition having antitumor effect and containing DNA

encoding antigenic protein

Patent: JP 2002528519-A 8 03-SEP-2002;

MENARINI RICERCHE SPA

COMMENT OS Homo sapiens (human)

PN JP 2002528519-A/8

PD 03-SEP-2002

PF 18-OCT-1999 JP 2000579265

PR 30-OCT-1998 IT MI98A002330

PI DINO PALLENTE,ANNA MARIA D MASSIMO,RITA DESANTIS PC

A61K38/00,A61K35/76,A61K39/00,A61K48/00,A61P35/00,C12N15/09, PC

A61K37/02,

PC C12N15/00

CC Medicinal composition having antitumor effect and containing

CC DNA encoding

CC antigenic protein

CC Key Location/Qualifiers

FT source 1..891

FT Location/Qualifiers

1..891

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 57.3%; Score 34.4; DB 6; Length 891;

Best Local Similarity 73.3%; Pred. No. 0.49;

Matches 44; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGGGGTCAACAAGCCGCCAGACACTGCACCTGCGCCA 60
|||||
Db 622 GGAAGTACTGCTCCACCAACACACACGGTGTACCTCGGCTCCGATACCAAGCCGCCCA 681

RESULT 6

AX180276/c

LOCUS AX180276 132 bp DNA linear PAT 06-AUG-2001

DEFINITION Sequence 34 from Patent WO0146228.

ACCESSION AX180276

VERSION AX180276.1 GI:15132247

KEYWORDS

SOURCE synthetic construct

ORGANISM synthetic construct

REFERENCE 1

AUTHORS Crowe,J.S. and Ellis,J.H.

TITLE Nucleic acid vaccination

JOURNAL

Patent: WO 0146228-A 34 28-JUN-2001;

GLAXO GROUP LIMITED (GB)

FEATURES Location/Qualifiers

1..132

/organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32630"

/note="Oligonucleotide"

ORIGIN

Query Match 54.7%; Score 32.8; DB 6; Length 132;
Best Local Similarity 71.7%; Pred. No. 2.1;
Matches 43; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGGGGTCAACAAGCCGCCAGACACTGCACCTGCGCCA 60
|||||
Db 111 GGAGCTGCTCCACCTGCACACGGCGGTGTGACACACCTGACACACGTCGCGGCCCA 52

RESULT 7

AX180275

LOCUS AX180275 144 bp DNA linear PAT 06-AUG-2001

DEFINITION Sequence 33 from Patent WO0146228.

ACCESSION AX180275

VERSION AX180275.1 GI:15132246

KEYWORDS

SOURCE synthetic construct

ORGANISM synthetic construct

REFERENCE 1

AUTHORS Crowe,J.S. and Ellis,J.H.

TITLE Nucleic acid vaccination

JOURNAL Patent: WO 0146228-A 33 28-JUN-2001;

GLAXO GROUP LIMITED (GB)

FEATURES Location/Qualifiers

1..144

/organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32630"

/note="Oligonucleotide"

ORIGIN

Query Match 54.7%; Score 32.8; DB 6; Length 144;

Best Local Similarity 71.7%; Pred. No. 2.1;

Matches 43; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGGGGTCAACAAGCCGCCAGACACTGCACCTGCGCCA 60
|||||
Db 28 GGAGCTGCTCCACCTGCACACGGCGGTGTGACACACCTGACACACGTCGCGGCCCA 87

RESULT 8

AX703425

LOCUS AX703425 93 bp DNA linear PAT 03-APR-2003

DEFINITION Sequence 3 from Patent WO02086505.

ACCESSION AX703425

VERSION AX703425.1 GI:29538416

KEYWORDS

SOURCE synthetic construct

ORGANISM synthetic construct

REFERENCE 1

AUTHORS Benson,R.S.

TITLE Intracellular analysis

JOURNAL Patent: WO 02086505-A 3 31-OCT-2002;

THE VICTORIA UNIVERSITY OF MANCHESTER (GB)

FEATURES Location/Qualifiers

1..93

/organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32630"

/note="Artificial epitope construct"

ORIGIN

ORGANISM synthetic construct

artificial sequences.

REFERENCE

1 Benson, R.S.

TITLE Intracellular analysis

JOURNAL Patent: WO 02086505-A 13 31-OCT-2002;

THE VICTORIA UNIVERSITY OF MANCHESTER (GB)

FEATURES location/Qualifiers

source 1..6877

/organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32630"

/note="Engineered construct"

ORIGIN

Query Match 52.3%; Score 31.4; DB 6; Length 6877;

Best Local Similarity 85.4%; Pred. No. 4;

Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGCGGCTCAACAAGCGGCC 41

Db 701 GGATCTACCGCTCCTCCTGCCACGAGTCAACAAGCGCAC 741

RESULT 14

AX180282/c

LOCUS AX180282 132 bp DNA linear PAT 06-AUG-2001

DEFINITION Sequence 40 from Patent WO0146228.

ACCESSION AX180282

VERSION AX180282.1 GI:15132253

KEYWORDS

SOURCE synthetic construct

ORGANISM synthetic construct

artificial sequences.

REFERENCE

1

AUTHORS Crowe, J.S. and Ellis, J.H.

TITLE Nucleic acid vaccination

JOURNAL Patent: WO 0146228-A 40 28-JUN-2001;

GLAXO GROUP LIMITED (GB)

FEATURES location/Qualifiers

source 1..132

/organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32630"

/note="Oligonucleotide"

ORIGIN

Query Match 52.0%; Score 31.2; DB 6; Length 132;

Best Local Similarity 70.0%; Pred. No. 7.2;

Matches 42; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGCGGCTCAACAAGCGGCCGAGACCTGACCTGGCCCA 60

Db 111 GGAGCTGTCGCTCCACCTGCACACGCGGCTGTTGACAGCACTGACACCGTCCGGCGCCA 52

RESULT 15

AX180281

LOCUS AX180281 144 bp DNA linear PAT 06-AUG-2001

DEFINITION Sequence 39 from Patent WO0146228.

ACCESSION AX180281

VERSION AX180281.1 GI:15132252

KEYWORDS

SOURCE synthetic construct

ORGANISM synthetic construct

artificial sequences.

REFERENCE

1

AUTHORS Crowe, J.S. and Ellis, J.H.

TITLE Nucleic acid vaccination

JOURNAL Patent: WO 0146228-A 39 28-JUN-2001;

GLAXO GROUP LIMITED (GB)

FEATURES location/Qualifiers

source 1..144

ORIGIN

/organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32630"

/note="Oligonucleotide"

Query Match 52.0%; Score 31.2; DB 6; Length 144;

Best Local Similarity 70.0%; Pred. No. 7.2;

Matches 42; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGCGGCTCAACAAGCGGCCGAGACCTGACCTGGCCCA 60

Db 28 GGAGCTGTCGCTCCACCTGCACACGCGGCTGTTGACAGCACTGACACCGTCCGGCGCCA 87

Search completed: January 15, 2005, 22:36:53

Job time : 722.8 secs

This Page Blank (uspto)

Db 1 GGCAGCACCAGCCGCCGACACAGGGGTCAAGCGGCCAGACACTGCACCTGCGCCA 60

RESULT 2

US-10-057-136-9
; Sequence 9, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUEFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-9

Query Match
Best Local Similarity 86.7%; Score 52; DB 14; Length 60;
Matches 55; Conservative 0; Pred. No. 8.9e-09; Mismatches 5; Indels 0; Gaps 0;

QY 1 GGCAGCACCAGCCGCCGACACAGGGGTCAAGCGGCCAGACACTGCACCTGCGCCA 60
Db 1 GGAAGTACCGCTCCACCTGCACACAGGGGTCAAGCGGCCAGACACTGCACCTGCGCCA 60

RESULT 3

US-10-406-317-41
; Sequence 41, Application US/10406317
; Publication No. US20040019195A1
; GENERAL INFORMATION:
; APPLICANT: Schlom, Jeffrey;
; APPLICANT: Hodge, James;
; APPLICANT: Panicali, Dennis
; TITLE OF INVENTION: A recombinant vector expressing multiple constitutulatory
; FILE REFERENCE: 38163-0189
; CURRENT APPLICATION NUMBER: US/10/406,317
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US/09/856,988
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: PCT/US99/26866
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/111,582
; PRIOR FILING DATE: 1998-12-09
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: VECTOR
US-10-406-317-41

Query Match
Best Local Similarity 86.7%; Score 52; DB 16; Length 2297;
Matches 91.7%; Pred. No. 5.8e-09;

Matches 55; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GGCAGCACCAGCCGCCGACACAGGGGTCAAGCGGCCAGACACTGCACCTGCGCCA 60
Db 706 GGAAGTACCGCTCCACCTGCACACAGGGGTCAAGCGGCCAGACACTGCACCTGCGCCA 765

RESULT 4

US-10-057-136-12
; Sequence 12, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUEFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-12

Query Match
Best Local Similarity 61.0%; Score 36.6; DB 14; Length 60;
Matches 45; Conservative 0; Pred. No. 0.0021; Mismatches 14; Indels 0; Gaps 0;

QY 1 GGCAGCACCAGCCGCCGACACAGGGGTCAAGCGGCCAGACACTGCACCTGCGGCC 59
Db 1 GGTAGTACAGCGCCACCGCCGACATGCGCTCAAGCGCTCCGGATTACAGAGCCGCGCC 59

RESULT 5

US-10-057-136-4
; Sequence 4, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUEFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-4

Query Match 55.7%; Score 33.4; DB 14; Length 60;
Best Local Similarity 72.9%; Pred. No. 0.027;
Matches 43; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 GGACAGACCGCAGCCGCGCACACGGGTTCACAAGCGCGCCAGACACTGACCTGCGCC 59
1 GGACAGTACTGCACCGCAGCATGCGTAAACATCAGACCTGTATACAGACCTGCACC 59

RESULT 6

US-10-057-136-5
; Sequence 5, Application US/10057136
; Publication No. US20030021770A1

; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-5

Query Match 53.0%; Score 31.8; DB 14; Length 60;
Best Local Similarity 71.2%; Pred. No. 0.099;
Matches 42; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 GGACAGACCGCAGCCGCGCACACGGGTTCACAAGCGCGCCAGACACTGACCTGCGCC 59
1 GGATCCACCGCGCGCTGCGCAGGAGTGACGTGCGCGCCGACACGCGCCCGCTCC 59

RESULT 7

US-10-057-136-11
; Sequence 11, Application US/10057136
; Publication No. US20030021770A1

; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 60

; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-11

Query Match 51.0%; Score 30.6; DB 14; Length 60;
Best Local Similarity 73.6%; Pred. No. 0.26;
Matches 39; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 7 ACCGACCGCCCGCAGACGGGTTCACAAGCGCGCCAGACACTGACCTGCGCC 59
7 ACGGCACTTCAGCAGCAGGAGTGACGTCTGCACCCGACACCCGTCACGCTCC 59

RESULT 8

US-10-057-136-2
; Sequence 2, Application US/10057136
; Publication No. US20030021770A1

; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-2

Query Match 50.3%; Score 30.2; DB 14; Length 60;
Best Local Similarity 69.5%; Pred. No. 0.36;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGACAGACCGCAGCCGCGCACACGGGTTCACAAGCGCGCCAGACACTGACCTGCGCC 59
1 GGCTCCACCGCCCCCGCAGCCAGGAGTGACCTGCGCGCCCGGACACGAGCGCGCCC 59

RESULT 9

US-10-057-136-13
; Sequence 13, Application US/10057136
; Publication No. US20030021770A1

; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; NUMBER OF SEQ ID NOS: 20

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 78
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-13
```

```
Query Match
Best Local Similarity 50.3%; Score 30.2; DB 14; Length 78;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
```

```
QY 1 GGCAGCACCAGCCGCGCCGACACAGGGGTCAACAGCGCGCCAGACTCGACCTGCGCC 59
Db 1 GGCTCCACCGCAGCCCGCCAGCCCGACGCTGTCACTCGGCCCCGGACACAGCGCGGCCCC 59
```

RESULT 10

```
US-10-447-839A-75/c
; Sequence 75, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Kufe, Donald W.
; APPLICANT: Kharbanda, Surender
; APPLICANT: Weitman, Steven D.
; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT APPLICATION NUMBER: US/10/447,839A
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 75
; LENGTH: 1424
; TYPE: RNA
; ORGANISM: ARTIFICIAL
; FEATURE:
; OTHER INFORMATION: Synthesized Sequence
US-10-447-839A-75
```

```
Query Match
Best Local Similarity 50.3%; Score 30.2; DB 16; Length 1424;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
```

```
QY 1 GGCAGCACCAGCCGCGCCGACACAGGGGTCAACAGCGCGCCAGACTCGACCTGCGCC 59
Db 1041 GGCTCCACCGCAGCCCGCCAGCCCGACGCTGTCACTCGGCCCCGGACACAGCGCGGCCCC 983
```

RESULT 11

```
US-10-447-839A-20
; Sequence 20, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Kufe, Donald W.
; APPLICANT: Kharbanda, Surender
; APPLICANT: Weitman, Steven D.
; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT APPLICATION NUMBER: US/10/447,839A
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.2
```

```
; SEQ ID NO 20
; LENGTH: 1428
; TYPE: RNA
; ORGANISM: RNA
US-10-447-839A-20
```

```
Query Match
Best Local Similarity 50.3%; Score 30.2; DB 16; Length 1428;
Matches 40; Conservative 1; Mismatches 18; Indels 0; Gaps 0;
```

```
QY 1 GGCAGCACCAGCCGCGCCGACACAGGGGTCAACAGCGCGCCAGACTCGACCTGCGCC 59
Db 385 GGCTCCACCGCAGCCCGCCAGCCCGACGCTGTCACTCGGCCCCGGACACAGCGCGGCCCC 443
```

RESULT 12

```
US-10-057-136-19
; Sequence 19, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
```

```
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: Kufe, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 1527
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-19
```

```
Query Match
Best Local Similarity 50.3%; Score 30.2; DB 14; Length 1527;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
```

```
QY 1 GGCAGCACCAGCCGCGCCGACACAGGGGTCAACAGCGCGCCAGACTCGACCTGCGCC 59
Db 226 GGCTCCACCGCAGCCCGCCAGCCCGACGCTGTCACTCGGCCCCGGACACAGCGCGGCCCC 284
```

RESULT 13

```
US-10-447-839A-19
; Sequence 19, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Kufe, Donald W.
; APPLICANT: Kharbanda, Surender
; APPLICANT: Weitman, Steven D.
; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT APPLICATION NUMBER: US/10/447,839A
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.2
```


SEQ ID NO 19
LENGTH: 1799
TYPE: RNA
ORGANISM: RNA
US-10-447-839A-19

Query Match 50.3%; Score 30.2; DB 16; Length 1799;
Best Local Similarity 67.8%; Pred. No. 0.24;
Matches 40; Conservative 1; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGCAGACCGCACCGCCCGCACACGGGGTCAACAAGCGCGCCAGACACTCGACCTGCGCC 59
Db 457 GGCTCCACCGCGCCCGCCCGCACCGGTGTCACTCGCGCCCGGACACCAAGCGCGCCCGCC 511

RESULT 14
US-09-964-824A-573
Sequence 573, Application US/09964824A
Patent No. US20020102531A1
GENERAL INFORMATION:

APPLICANT: Horrigan, Stephen
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
FILE REFERENCE: 689290-73
CURRENT APPLICATION NUMBER: US/09/964, 824A
CURRENT FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: US/60/236, 033
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US/60/236, 032
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US/60/236, 028
NUMBER OF SEQ ID NOS: 583
SOFTWARE: PatentIn version 3.0
SEQ ID NO 573
LENGTH: 1804
TYPE: DNA
ORGANISM: Homo sapiens
US-09-964-824A-573

Query Match 50.3%; Score 30.2; DB 9; Length 1804;
Best Local Similarity 69.5%; Pred. No. 0.24;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGCAGACCGCACCGCCCGCACACGGGGTCAACAAGCGCGCCAGACACTCGACCTGCGCC 59
Db 457 GGCTCCACCGCGCCCGCCCGCACCGGTGTCACTCGCGCCCGGACACCAAGCGCGCCCGCC 515

RESULT 15
US-10-029-517-17
Sequence 17, Application US/10029517
Publication No. US20030148969A1
GENERAL INFORMATION:
APPLICANT: Kenneth W. Dobie
APPLICANT: Susan J. Myers
TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
FILE REFERENCE: RTS-0352
CURRENT APPLICATION NUMBER: US/10/029, 517
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 107
SEQ ID NO 17
LENGTH: 1804
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (73) ... (1500)
US-10-029-517-17

Query Match 50.3%; Score 30.2; DB 15; Length 1804;
Best Local Similarity 69.5%; Pred. No. 0.24;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGCAGACCGCACCGCCCGCACACGGGGTCAACAAGCGCGCCAGACACTCGACCTGCGCC 59
Db 457 GGCTCCACCGCGCCCGCCCGCACCGGTGTCACTCGCGCCCGGACACCAAGCGCGCCCGCC 515

Search completed: January 16, 2005, 09:30:25
Job time : 183.4 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: January 15, 2005, 19:50:13 ; Search time 35.2 Seconds
(without alignments)
1211.572 Million cell updates/sec

Title: US-10-057-136-8

Perfect score: 60
Sequence: 1 GGCAGCACCAGCCGCGCCGCGC.....CAGCAGCTGACCTGCGCCA 60

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfillseq1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30.2	50.3	1804	4 US-10-029-517-17	Sequence 17, Appl
2	30.2	50.3	8186	4 US-10-029-517-19	Sequence 19, Appl
3	28.8	48.0	518	4 US-10-029-517-101	Sequence 101, Appl
4	28.6	47.7	572	4 US-10-029-517-18	Sequence 18, Appl
5	28.6	47.7	1721	4 US-10-029-517-3	Sequence 3, Appl
6	27.4	45.7	60	4 US-09-475-947A-246	Sequence 246, App
7	27.2	45.3	3343	4 US-10-029-517-102	Sequence 102, App
8	26	43.3	6192	2 US-08-479-537A-1	Sequence 1, Appl
9	26	43.3	6192	3 US-09-083-116-1	Sequence 1, Appl
10	26	43.3	6192	3 US-09-134-916A-1	Sequence 1, Appl
11	26	43.3	6449	2 US-08-479-537A-4	Sequence 4, Appl
12	26	43.3	6449	3 US-09-083-116-4	Sequence 4, Appl
13	26	43.3	6449	3 US-09-134-916A-4	Sequence 4, Appl
14	25.6	42.7	981	4 US-10-029-517-16	Sequence 16, Appl
15	25	41.7	4041	4 US-09-569-611C-1	Sequence 1, Appl
16	24.8	41.3	1174	3 US-09-034-985-1	Sequence 1, Appl
17	24.8	41.3	1417	4 US-09-535-008-33	Sequence 33, Appl
18	24.4	40.7	846	3 US-09-455-960-5	Sequence 5, Appl
19	24.4	40.7	846	4 US-10-051-325-5	Sequence 2, Appl
20	24.4	40.7	2500	3 US-09-455-960-2	Sequence 5, Appl
21	24.4	40.7	2500	4 US-10-051-325-2	Sequence 2, Appl
22	24.4	40.7	2748	4 US-09-799-451-598	Sequence 598, App
23	24.2	40.3	1392	4 US-09-489-039A-5663	Sequence 5663, App
24	24	40.0	2900	1 US-08-034-650-9	Sequence 9, Appl
25	23.8	39.7	633	4 US-09-252-991A-3966	Sequence 3966, Ap
26	23.8	39.7	1899	4 US-09-252-991A-4034	Sequence 4034, Ap

28	23.8	39.7	1920	4 US-09-252-991A-4000	Sequence 4000, Ap
29	23.4	39.0	154	4 US-09-513-999C-26161	Sequence 26161, A
30	23.4	39.0	888	4 US-09-311-021-155	Sequence 155, App
31	23.4	39.0	920	4 US-09-620-312D-132	Sequence 132, App
32	23.4	39.0	4403765	3 US-09-103-840A-2	Sequence 2, Appl
33	23.4	39.0	4411529	3 US-09-103-840A-1	Sequence 1, Appl
34	23.2	38.7	3957	1 US-07-689-008-5	Sequence 5, Appl
35	23.2	38.7	9540	1 US-07-689-008-1	Sequence 1, Appl
36	23	38.3	274	4 US-09-621-976-16450	Sequence 16450, A
37	23	38.3	425	4 US-08-956-171E-3693	Sequence 3693, Ap
38	23	38.3	425	4 US-08-781-986A-3693	Sequence 3693, Ap
39	23	38.3	918	4 US-09-252-991A-9303	Sequence 9266, Ap
40	23	38.3	1428	4 US-09-252-991A-9266	Sequence 9274, Ap
41	23	38.3	1545	4 US-09-252-991A-9274	Sequence 135, App
42	22.8	38.0	355	4 US-09-370-838-135	Sequence 135, App
43	22.8	38.0	355	4 US-09-854-133-135	Sequence 135, App
44	22.8	38.0	507	4 US-08-956-171E-3895	Sequence 3895, Ap
45	22.8	38.0	507	4 US-08-781-986A-3895	Sequence 3895, Ap

ALIGNMENTS

RESULT 1
US-10-029-517-17
; Sequence 17, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Susan J. Myers
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 17
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (73)...(1500)
; US-10-029-517-17

Query Match 50.3%; Score 30.2; DB 4; Length 1804;
Best Local Similarity 69.5%; Pred. No. 0.26;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGCAGCACCAGCCGCGCCGACAGCGGTCACAMAGCGCCAGACACTGACCTGCGCC 59
Db 457 GGCTCCAGCCGCCGCCAGCCAGCGTGTACCTCGGCCCGGACACAGCGCGCCC 515

RESULT 2
US-10-029-517-19
; Sequence 19, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 19
; LENGTH: 8186
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 6899
; OTHER INFORMATION: unknown

```
; NAME/KEY: unsure
; LOCATION: 7155
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7184
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7957
; OTHER INFORMATION: unknown
; NAME/KEY: intron
; LOCATION: (2997)...(3498)
; OTHER INFORMATION: intron 1
; NAME/KEY: intron:exon junction
; LOCATION: (3498)...(3499)
; OTHER INFORMATION: intron 1:exon 2
; NAME/KEY: exon
; LOCATION: (3508)...(3599)
; OTHER INFORMATION: exon 2d
; NAME/KEY: exon:intron junction
; LOCATION: (3982)...(3983)
; OTHER INFORMATION: exon 2a:intron 2a
; NAME/KEY: intron:exon junction
; LOCATION: (4205)...(4206)
; OTHER INFORMATION: intron 2c:exon 3c
; NAME/KEY: intron:exon junction
; LOCATION: (4259)...(4260)
; OTHER INFORMATION: intron 2d:exon 3d
; NAME/KEY: exon
; LOCATION: (4260)...(4328)
; OTHER INFORMATION: exon 3d
; NAME/KEY: intron:exon junction
; LOCATION: (4632)...(4633)
; OTHER INFORMATION: intron 3:exon 4
; NAME/KEY: exon
; LOCATION: (4914)...(5035)
; OTHER INFORMATION: exon 5
; NAME/KEY: intron
; LOCATION: (5266)...(6293)
; OTHER INFORMATION: intron 6
US-10-029-517-19
```

Query Match 50.3%; Score 30.2; DB 4; Length 8186;
Best Local Similarity 69.5%; Pred. No. 0.3;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```
QY 1 GGCAGCACCAGCAGCCCGCCGACACAGGGGTCAACAAGCGCCGACACTCGACCTGCGCC 59
Db 3825 GGCTCCACCGCCCCCCCCAGCCCGGTCACCTCGGCCCCGACACCAAGCGCGCCCC 3883
```

RESULT 3

```
US-10-029-517-101
; Sequence 101, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 101
; LENGTH: 518
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-10-029-517-101
```

Query Match 48.0%; Score 28.8; DB 4; Length 518;
Best Local Similarity 69.6%; Pred. No. 0.68;
Matches 39; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```
QY 1 GGCAGCACCAGCAGCCCGCCGACACAGGGGTCAACAAGCGCCGACACTCGACCTGCC 56
Db 462 GGCTCCACCGCCCCCCCCAGCCCGGTCACCTCGGCCCCGACACCAAGCGCGCC 517
```

RESULT 4

```
US-10-029-517-18
; Sequence 18, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 18
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (67)...(572)
US-10-029-517-18
```

Query Match 47.7%; Score 28.6; DB 4; Length 572;
Best Local Similarity 67.8%; Pred. No. 0.8;
Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```
QY 1 GGCAGCACCAGCAGCCCGCCGACACAGGGGTCAACAAGCGCCGACACTCGACCTGCGCC 59
Db 478 GGCTCCACCGCCCCCCCCAGCCCGGTCACCTCGGCCCCGACACCAAGCGCGCCCC 536
```

RESULT 5

```
US-10-029-517-3
; Sequence 3, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 3
; LENGTH: 1721
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (58)...(1605)
US-10-029-517-3
```

Query Match 47.7%; Score 28.6; DB 4; Length 1721;
Best Local Similarity 67.8%; Pred. No. 0.88;
Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```
QY 1 GGCAGCACCAGCAGCCCGCCGACACAGGGGTCAACAAGCGCCGACACTCGACCTGCGCC 59
Db 442 GGCTCCACCGCCCCCCCCAGCCCGGTCACCTCGGCCCCGACACCAAGCGCGCCCC 500
```

RESULT 6

```
US-09-475-947A-246
; Sequence 246, Application US/09475947A
; Patent No. 6472154
; GENERAL INFORMATION:
; APPLICANT: Garner, Harold R.
; APPLICANT: Wren, Jonathan D.
; APPLICANT: Minna, John D.
```


US-09-083-116-1
; Sequence 1, Application US/09083116
; Patent No. 6203795
; GENERAL INFORMATION:
; APPLICANT: CHAMBERON, Pierre
; APPLICANT: KIENY, Marie-Paule
; APPLICANT: LATHE, Richard
; APPLICANT: HAREUVENI, Mara
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
; TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/083,116
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,537
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR91/00835
; FILING DATE: 23-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/039,320
; FILING DATE: 04-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,576
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 017753-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6192 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 58..120
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 439..5239
; OTHER INFORMATION: /note= "The nucleotides spanning
; 439-5239 constitute a repeated region wherein the repeat is 6
; OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
; OTHER INFORMATION: The number of such repeats varies from 1 to 80."
; NAME/KEY: mat_peptide
; LOCATION: 121..6166
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 457
; OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG, and Ala = GCT, GCC, GCA, or GCG."
; FEATURE:

NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG, and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG, and Ala = GCT, GCC, GCA, or GCG."
US-09-083-116-1
Query Match 43.3%; Score 26; DB 3; Length 6192;
Best Local Similarity 59.3%; Pred. No. 7;
Matches 35; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 1 GGACGACCGGACCGCCCGACACGCGGTCAAGCGCGCCAGACACTGACCTGCGCC 59
DB 442 GGCTCCACCGCCCGCCNNMGCCGACGCTGTCACCTCGGCCCGGACNNNAGGCCGNNCC 500
RESULT 10
US-09-134-916A-1
; Sequence 1, Application US/09134916A
; Patent No. 6328956
; GENERAL INFORMATION:
; APPLICANT: CHAMBERON, Pierre
; APPLICANT: KIENY, Marie-Paule
; APPLICANT: LATHE, Richard
; APPLICANT: HAREUVENI, Mara
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
; TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/134,916A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,537
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: FR 90/13101
; FILING DATE: 23-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR91/00835
; FILING DATE: 23-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/039,320
; FILING DATE: 04-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,576
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 017753-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

```

; LENGTH: 6192 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 58..120
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 439..5239
; OTHER INFORMATION: /note= "The nucleotides spanning
; OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
; OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
; OTHER INFORMATION: The number of such repeats varies from 1 to 80."
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 121..6166
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 457
; OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 487
; OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
; OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
; OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 496
; OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; US-09-134-916A-1

Query Match          43.3%; Score 26; DB 3; Length 6192;
Best Local Similarity 59.3%; Pred. No. 7;
Matches 35; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY      1  GGACAGACCGACCGCCGCGACACAGCGGTCACAAGCGGCCAGACACTGACCTGCGCC 59
Db      442 GGCTCCACCGCCGCCNNNGCCACGCGTGTCACTCGGCCCGGACNNNAGGCCGNNNC 500

RESULT 11
US-08-479-537A-4
; Sequence 4, Application US/08479537A
; Patent No. 5861381
; GENERAL INFORMATION:
; APPLICANT: CHAMON, Piere
; APPLICANT: KIENY, Marie-Paule
; APPLICANT: LATHE, Richard
; APPLICANT: HAREUVENI, Mara
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
; TITLE OF INVENTION: TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,537A
```

```

; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 90/13101
; FILING DATE: 23-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR91/00835
; FILING DATE: 23-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/039,320
; FILING DATE: 04-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,576
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 017753-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6449 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 58..120
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 439..5239
; OTHER INFORMATION: /note= "The nucleotides spanning
; OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is
; OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixe
; OTHER INFORMATION: The number of such repeats varies from 1 to 80."
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 121..5661
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 457
; OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 487
; OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
; OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
; OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 496
; OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; US-08-479-537A-4

Query Match          43.3%; Score 26; DB 2; Length 6449;
Best Local Similarity 59.3%; Pred. No. 7;
Matches 35; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY      1  GGACAGACCGACCGCCGCGACACAGCGGTCACAAGCGGCCAGACACTGACCTGCGCC 59
Db      442 GGCTCCACCGCCGCCNNNGCCACGCGTGTCACTCGGCCCGGACNNNAGGCCGNNNC 500

RESULT 12
US-09-083-116-4
; Sequence 4, Application US/09083116
```


Patent No. 6203795
GENERAL INFORMATION:
APPLICANT: CHAMBERN, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREUVENI, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/083,116
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,537
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6449 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note= "The nucleotides spanning
439-5239 constitute a repeated region wherein the repeat is 6
nucleotides and encodes 20 amino acids, 17 of which are fixed
OTHER INFORMATION: The number of such repeats varies from 1 to 80."
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 121..5661
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG, and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487

OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG, and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG, and Ala = GCT, GCC, GCA, or GCG."
US-09-083-116-4
Query Match 43.3%; Score 26; DB 3; Length 6449;
Best Local Similarity 59.3%; Pred. No. 7;
Matches 35; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
Oy 1 GGACGACCGCAGCCCGCAGCAGGGGTCAAGCGCCGACACTCGACCTGCGCC 59
Db 442 GGCTCCACCGCCGCCNNNGCCACGCGTGTCACTCGGCCCGGACNNAGGCCGNNCC 500
RESULT 13
US-09-134-916A-4
Sequence 4, Application US/09134916A
Patent No. 6328956
GENERAL INFORMATION:
APPLICANT: CHAMBERN, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREUVENI, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/134,916A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/479,537
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: FR 90/13101
FILING DATE: 23-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6449 base pairs
TYPE: nucleic acid

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 58..120
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 439..5239
; OTHER INFORMATION: /note= "The nucleotides spanning
; OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
; OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
; OTHER INFORMATION: The number of such repeats varies from 1 to 80."
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 121..5661
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 457
; OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 487
; OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
; OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
; OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 496
; OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; US-09-134-916A-4

Query Match          43.3%; Score 26; DB 3; Length 6449;
Best Local Similarity 59.3%; Pred. No. 7;
Matches 35; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY      1 GGACAGACCGCACCAGCGCCGACACAGCGGCTCACAAGCGCGCCAGACACTGCACCTGCC 59
         ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      442 GGCTCCACCGCCCCCANNNGCCACCGGTGTCACTCGGCCCGGACNNNAGGCCGNNNCC 500

RESULT 14
US-10-029-517-16
; Sequence 16, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 16
; LENGTH: 981
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon:exon junction
; LOCATION: (464)..(465)
; OTHER INFORMATION: exon 3b:exon 4
; US-10-029-517-16

Query Match          42.7%; Score 25.6; DB 4; Length 981;
Best Local Similarity 66.1%; Pred. No. 8.1;
Matches 37; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
```

```
QY      1 GGACAGACCGCACCAGCGCCGACACAGCGGCTCACAAGCGCGCCAGACACTGCACCTGC 56
         ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db      21 GGCTCCACCGCCCCCAGCCCATGTGTACCTCGGCCCGGACACAGCCCCGC 76

RESULT 15
US-09-569-611C-1/c
; Sequence 1, Application US/09569611C
; Patent No. 6720182
; GENERAL INFORMATION:
; APPLICANT: SAVITZKY et al.
; TITLE OF INVENTION: ALTERNATIVE SPLICING VARIANTS
; FILE REFERENCE: 2786-0151P
; CURRENT APPLICATION NUMBER: US/09/569,611C
; CURRENT FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4041
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-569-611C-1

Query Match          41.7%; Score 25; DB 4; Length 4041;
Best Local Similarity 64.9%; Pred. No. 14;
Matches 37; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
```

```
QY      1 GGACAGACCGCACCAGCGCCGACACAGCGGCTCACAAGCGCGCCAGACACTGCACCTGC 57
         ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      2244 GGCGGTACAGACCGGCTGCTCAGCGGGCACCTTGTCTCCTGATCTCGAATCCG 2188
```

Search completed: January 16, 2005, 03:01:42
Job time : 40.2 secs

This Page Blank (uspto)

TITLE	21C Frontier Korean EST Project 2001
JOURNAL	Unpublished (2002)
COMMENT	Contact: Kim YS

Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 14 row: A column: 06
High quality sequence stop: 604.

FEATURES	Location/Qualifiers
source	1. .604

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S21SNU520-14-A06"
/sex="F"
/tissue_type="Stomach"
/cell_type="Floating aggregates"
/cell_line="SNU-520"
/lab_host="Top10F"
/clone_lib="S21SNU520"
/notes="Organ: Stomach; Vector: pTZ19RP1; Site_1: EcoRI; Site_2: NotI; The poly (A) + RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoRI I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

```

ORIGIN

```

Query Match      53.0%; Score 31.8; DB 4; Length 604;
Best Local Similarity 71.2%; Pred. No. 9.2;
Matches 42; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```

[illegible]

RESULT 4	BU542454	877 bp	mRNA	linear	EST 13-SEP-2002
LOCUS	AGENCOURT_10322173	NIH_MGC_40	Homo sapiens	CDNA clone	IMAGE:65743322
DEFINITION	5', mRNA sequence.				

ACCESSION	BU542454	
VERSION	BU542454.1	GI:22852937
KEYWORDS	EST.	
SOURCE	Homo sapiens	
ORGANISM	Homo sapiens (human)	

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 877)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

CDNA Library Preparation: Rubin laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be

```

found through the I.M.A.G.E: Consortium/LLNL at
http://image.llnl.gov
Plate: LLCM2769 row: h column: 02
High quality sequence stop: 760.
location/Qualifiers
1..877
FEATURES
source
```

FEATURES

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6574322"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/notes="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH MGC Library."

```

ORIGIN

Query Match	53.0%	Score 31.8;	DB 5;	length 877;
Best Local Similarity	71.2%;	Pred. No. 9;		
Matches 42; Conservative	0;	Mismatches 17;	Indels 0;	Gaps 0;

Oy 1 GGAGCAGCCGCA CCGCCCCGCACA CGGGGTCA CAAGCGCGCCAGACACTGCCTGCGGCC 59
||| |||||||| | | | | | | | | | | | | | | | |
Db 215 GGGTCACCGCACCCCAGCCGCCA CGGTGTCACTTGCGCCCCGGGACACCAAGCGCGGCCCC 27

RESULT 5	
CA489836	
LOCUS	959 bp
DEFINITION	AGENCOURT_10810668 MABCL Homo sapiens CDNA clone IMAGE:6722324 5', mRNA sequence.
	linear EST 14-NOV-2002

ACCESSION	CA489836
VERSION	CA489836.1
KEYWORDS	GI:24952627
SOURCE	Homo sapiens (human) EST.

ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 959)
REFERENCE	NIH-MGC http://mgc.nci.nih.gov/ .
AUTHORS	National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE	Unpublished (1999)
JOURNAL	Contact: Robert Strausberg, Ph.D.
COMMENT	

Tissue Procurement: Kristi A. Egland, Ira Pastan
 cDNA Library Preparation: Invitrogen Corp
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLNL4284 row: n column: 20
 High quality sequence start: 31
 High quality sequence stop: 446.

FEATURES
SOURCE

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6722324"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231,
hTERT-HMEL, LNCaP"
/lab_host="EMDH10B"
/clone_idb="MAPcl"
/notes="vector: pCMV-SPORT6; site 1: EcoRV; site 2: Not I
Subtracted with brain, liver, lung, kidney and muscle.
directionally cloned. Priming method: oligo-dT. Average

```

Insert size: 1800 bp. Library amplification: 26,000 fold.
Kristi A. Egland, James J. Vincent, Robert Strausberg,
Bungkook Lee & Ira Pastan: Discovery of new breast
cancer genes encoding membrane and secreted proteins.
Manuscript submitted."

ORIGIN
Query Match 53.0%; Score 31.8; DB 6; Length 959;
Best Local Similarity 71.2%; Pred. No. 9;
Matches 42; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 1 GGACAGCCGACCGCCGACACGGGGTCAACAGCCGCCAGACACTCGACCTGCGCC 59
Db 469 GGCTCCACCGCACCCGCCAGCCCGGTGTCACTCTCGCCCGGACACAGCGCGGCC 527

RESULT 6
AUI92332/c 499 bp mRNA linear EST 14-OCT-2003
LOCUS AUI92332 Porphyrta yezeensis TU-1 sporophytes Porphyrta yezeensis
DEFINITION cDNA clone PFL029d07_r 5', mRNA sequence.
ACCESSION AUI92332
VERSION AUI92332.1 GI:31930868
KEYWORDS EST.
SOURCE Porphyrta yezeensis
ORGANISM Porphyrta yezeensis

REFERENCE 1 (bases 1 to 499)
AUTHORS Asamizu,E., Nakajima,M., Kitade,Y., Saga,N., Nakamura,Y. and Tabata,S.
TITLE COMPARISON OF RNA EXPRESSION PROFILES BETWEEN THE TWO GENERATIONS OF PORPHYRTA YEZEENSIS (RHODOPHYTA), BASED ON EXPRESSED SEQUENCE TAG FREQUENCY ANALYSIS
JOURNAL J. Phycol. 39 (5), 923-930 (2003)
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
source 1.499
/organism="Porphyrta yezeensis"
/mol_type="mRNA"
/strain="TU-1"
/db_xref="taxon:2788"
/clone="PFL029d07_r"
/dev_stage="sporophytes"
/clone_lib="Porphyrta yezeensis TU-1 sporophytes"

ORIGIN
Query Match 52.3%; Score 31.4; DB 1; Length 499;
Best Local Similarity 77.6%; Pred. No. 12;
Matches 38; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 8 CCGCAGCCGCGACACGGGGTCAACAGCGCGCCAGACACTCGACCTGC 56
Db 295 CCGCAGCCGCGCGACCGGGGTAAACAAGAGACCGAGACGATCACCAGC 247

RESULT 7
AUI94847/c 520 bp mRNA linear EST 14-OCT-2003
LOCUS AUI94847 Porphyrta yezeensis TU-1 sporophytes Porphyrta yezeensis
DEFINITION cDNA clone PFL066b09_r 5', mRNA sequence.
ACCESSION AUI94847
VERSION AUI94847.1 GI:31935887
KEYWORDS EST.
SOURCE Porphyrta yezeensis
ORGANISM Porphyrta yezeensis

REFERENCE 1 (bases 1 to 520)
Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;
Porphyrta.

AUTHORS Asamizu,E., Nakajima,M., Kitade,Y., Saga,N., Nakamura,Y. and Tabata,S.
TITLE COMPARISON OF RNA EXPRESSION PROFILES BETWEEN THE TWO GENERATIONS OF PORPHYRTA YEZEENSIS (RHODOPHYTA), BASED ON EXPRESSED SEQUENCE TAG FREQUENCY ANALYSIS
JOURNAL J. Phycol. 39 (5), 923-930 (2003)
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
source 1.520
/organism="Porphyrta yezeensis"
/mol_type="mRNA"
/strain="TU-1"
/db_xref="taxon:2788"
/clone="PFL066b09_r"
/dev_stage="sporophytes"
/clone_lib="Porphyrta yezeensis TU-1 sporophytes"

ORIGIN
Query Match 52.3%; Score 31.4; DB 1; Length 520;
Best Local Similarity 77.6%; Pred. No. 12;
Matches 38; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 8 CCGCAGCCGCGACACGGGGTCAACAGCGCGCCAGACACTCGACCTGC 56
Db 209 CCGCAGCCGCGCGACCGGGGTAAACAAGAGACCGAGACGATCACCAGC 161

RESULT 8
AUI94863/c 530 bp mRNA linear EST 14-OCT-2003
LOCUS AUI94863 Porphyrta yezeensis TU-1 sporophytes Porphyrta yezeensis
DEFINITION cDNA clone PFL066d09_r 5', mRNA sequence.
ACCESSION AUI94863
VERSION AUI94863.1 GI:31935919
KEYWORDS EST.
SOURCE Porphyrta yezeensis
ORGANISM Porphyrta yezeensis

REFERENCE 1 (bases 1 to 530)
AUTHORS Asamizu,E., Nakajima,M., Kitade,Y., Saga,N., Nakamura,Y. and Tabata,S.
TITLE COMPARISON OF RNA EXPRESSION PROFILES BETWEEN THE TWO GENERATIONS OF PORPHYRTA YEZEENSIS (RHODOPHYTA), BASED ON EXPRESSED SEQUENCE TAG FREQUENCY ANALYSIS
JOURNAL J. Phycol. 39 (5), 923-930 (2003)
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
source 1.530
/organism="Porphyrta yezeensis"
/mol_type="mRNA"
/strain="TU-1"
/db_xref="taxon:2788"
/clone="PFL066d09_r"
/dev_stage="sporophytes"
/clone_lib="Porphyrta yezeensis TU-1 sporophytes"

ORIGIN
Query Match 52.3%; Score 31.4; DB 1; Length 530;
Best Local Similarity 77.6%; Pred. No. 12;
Matches 38; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 8 CCGCAGCCGCGACACGGGGTCAACAGCGCGCCAGACACTCGACCTGC 56
Db 209 CCGCAGCCGCGCGACCGGGGTAAACAAGAGACCGAGACGATCACCAGC 161

RESULT 9
LOCUS AU193189/c 552 bp mRNA linear EST 14-OCT-2003
DEFINITION AU193189 Porphyra yezoensis TU-1 sporophytes Porphyra yezoensis
cDNA clone PFL041f12_r 5', mRNA sequence.
ACCESSION AU193189
VERSION AU193189
KEYWORDS AU193189.1 GI:31932583
SOURCE EST.
ORGANISM Porphyra yezoensis
Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;
Porphyra.
REFERENCE 1 (bases 1 to 552)
AUTHORS Asamizu,E., Nakajima,M., Kitade,Y., Saga,N., Nakamura,Y. and
Tabata,S.
TITLE COMPARISON OF RNA EXPRESSION PROFILES BETWEEN THE TWO GENERATIONS
OF PORPHYRA YEZOENSIS (RHODOPHYTA), BASED ON EXPRESSED SEQUENCE TAG
FREQUENCY ANALYSIS
JOURNAL J. Phycol. 39 (5), 923-930 (2003)
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/
location/Qualifiers
FEATURES
Source 1..552
/organism="Porphyra yezoensis"
/mol_type="mRNA"
/strain="TU-1"
/db_xref="taxon:2788"
/clone="PFL041f12_r"
/dev_stage="sporophytes"
/clone_lib="Porphyra yezoensis TU-1 sporophytes"
ORIGIN
Query Match 52.3%; Score 31.4; DB 1; Length 552;
Best Local Similarity 77.6%; Pred. No. 12;
Matches 38; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 8 CCGCACC GCCGACACGGGGTCAACAAGCCGCCAGACACTGCACCTGC 56
|||||
Db 302 CCGCACC GCCGACACGGGGTCAACAAGAGACGACGACATCACCAGC 254
|||||
RESULT 10
LOCUS AI925867/c 330 bp mRNA linear EST 08-MAR-2000
DEFINITION AI925867
wo20604.x1 NCI CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2455879 3',
similar to contains element MSRI MSRI repetitive element ;, mRNA
sequence.
ACCESSION AI925867
VERSION AI925867
KEYWORDS AI925867.1 GI:5661831
SOURCE EST.
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 330)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www-bio.lnl.gov/bbrp/image/image.html
Insert length: 578 Std Error: 0.00
Seq primer: -40UP from Gibco

High quality sequence stop: 305.
FEATURES
Source Location/Qualifiers
1..330
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2455879"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/clone_lib="NCI CGAP Pan1"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.72 kb. Life Technologies catalog #: 11548-013"
ORIGIN
Query Match 50.3%; Score 30.2; DB 1; Length 330;
Best Local Similarity 69.5%; Pred. No. 29;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY 1 GGCAGACCGCACC GCCGACACGGGGTCAACAAGCCGCCAGACACTGCACCTGC 59
|||||
Db 316 GGCTCCACCGCCCCCAGCCACCGGTGTCACTGCGCCCGGACACGCGGCC 258
|||||
RESULT 11
LOCUS BU148487 1113 bp mRNA linear EST 03-SEP-2002
DEFINITION AGENCOURT 8670479 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6380642
5', mRNA sequence.
ACCESSION BU148487
VERSION BU148487.1 GI:22662019
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1113)
AUTHORS NIH-MGC http://mhc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LCM2569 row: j column: 03
High quality sequence stop: 235.
FEATURES
Source Location/Qualifiers
1..1113
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6380642"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
ORIGIN
Query Match 50.3%; Score 30.2; DB 5; Length 1113;
Best Local Similarity 69.5%; Pred. No. 28;

This Page Blank (uspto)

OM nucleic - nucleic search, using BW model

Title: US-10-057-136-8

Scoring table: IDENTITY_NUC

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing:	Minimum Match 0%
	Maximum Match 100%
	Listing first 45 summaries

```
Database :
1: _N_Geneseq_23Sep04:*
2: genesegn1980s:*
3: genesegn2000s:*
4: genesegn2001as:*
5: genesegn2001bs:*
6: genesegn2002as:*
7: genesegn2002bs:*
8: genesegn2003as:*
9: genesegn2003bs:*
10: genesegn2003cs:*
11: genesegn2003ds:*
12: genesegn2004s:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	60	100.0	60	2	AAV48321	Aav48321 Nucleotid
2	52	86.7	60	2	AAV48322	Aav48322 Nucleotid
3	36.6	61.0	60	2	AAV48325	Aav48325 Nucleotid
4	33.4	55.7	60	2	AAV48317	Aav48317 Nucleotid
5	31.8	53.0	60	2	AAV48318	Aav48318 Nucleotid
6	30.6	51.0	60	2	AAV48324	Aav48324 Nucleotid
7	30.2	50.3	60	2	AAV48316	Aav48316 Nucleotid
8	30.2	50.3	78	2	AAV48326	Aav48326 Nucleotid
9	30.2	50.3	309	1	AAN90579	Aan90579 PDF9.3 CD
10	30.2	50.3	1194	12	ADi57712	Adi57712 Human bre
11	30.2	50.3	1378	12	ADi57693	Adi57693 Human bre
12	30.2	50.3	1424	12	ADO23180	Ado23180 Antisense
13	30.2	50.3	1428	6	ABL60159	Abi60159 Human MUC
14	30.2	50.3	1428	12	ADO23125	Ado23125 Human MUC
15	30.2	50.3	1457	12	ADP32627	Adf32627 Plasmid J
16	30.2	50.3	1527	2	AAV48329	Aav48329 MiniMUC1
17	30.2	50.3	1572	5	AAS00585	Aas00585 Human MUC
18	30.2	50.3	1614	12	ADK70370	Adk70370 Respiratc
19	30.2	50.3	1630	12	ADi57708	Adi57708 Human bre
20	30.2	50.3	1634	12	ADi57689	Adi57689 Human bre
21	30.2	50.3	1712	12	ADi57686	Adi57686 Human bre

22	30.2	50.3	1738	12	AD157669	Ad157669 Human bre
23	30.2	50.3	1755	12	AD157673	Ad157673 Human bre
24	30.2	50.3	1774	12	AD543991	Ad543991 Plaemid J
25	30.2	50.3	1774	12	ADf32625	Adf32625 Plaemid J
26	30.2	50.3	1799	12	AD023124	Ad023124 Human MUC
27	30.2	50.3	1803	12	AD157699	Ad157699 Human bre
28	30.2	50.3	1804	6	AB167539	Ab167539 Thyroid c
29	30.2	50.3	1804	9	AAD56950	Aad56950 Human muc
30	30.2	50.3	1804	10	ADd14719	Ad d14719 Human src
31	30.2	50.3	1805	12	ADp13294	Adp13294 Renal cel
32	30.2	50.3	1805	12	AD028642	Ad028642 Human MUC
33	30.2	50.3	1808	12	AD157706	Ad157706 Human bre
34	30.2	50.3	1818	12	ADf32633	Adf32633 Plaemid J
35	30.2	50.3	1823	6	AB235228	Ab235228 Human gen
36	30.2	50.3	1823	12	AD157707	Ad157707 Human bre
37	30.2	50.3	1835	12	ADf32631	Adf32631 Plaemid J
38	30.2	50.3	1874	12	AD157688	Ad157688 Human bre
39	30.2	50.3	1882	12	AD157677	Ad157677 Human bre
40	30.2	50.3	1918	12	AD157672	Ad157672 Human bre
41	30.2	50.3	1930	12	AD157678	Ad157678 Human bre
42	30.2	50.3	1945	12	AD157676	Ad157676 Human bre
43	30.2	50.3	1949	12	AD157698	Ad157698 Human bre
44	30.2	50.3	1953	12	AD157668	Ad157668 Human bre
45	30.2	50.3	2045	12	AD157701	Ad157701 Human bre

ALIGNMENTS

CC	XX	RESULT 1
CC	XX	AAV48321
CC	AC	AAV48321 standard; DNA; 60 BP.
CC	XX	AAV48321;
CC	DT	20-NOV-1998 (first entry)
CC	DE	Nucleotide sequence encoding MUC1 tandem repeat unit R6.
CC	KW	ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
CC	XX	tumour; tumour-associated antigen.
CC	OS	Homo sapiens.
CC	PN	WO9837095-A2.
CC	PD	27-AUG-1998.
CC	PF	24-FEB-1998; 98WO-US003693.
CC	PR	24-FEB-1997; 97US-0038253P.
CC	PA	(THER-) THERION BIOLOGICS CORP.
CC	PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.
CC	PA	(DAND) DANA FARBER CANCER INST INC.
CC	PI	Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
CC	DR	WPL; 1998-467492/40.
CC	PT	New recombinant pox virus for tumour therapy - comprises DNA encoding an
CC	PT	immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
CC	PS	Disclosure; Page 11; 42pp; English.
CC	XX	The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC	XX	immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC	XX	(RPV). The RPV was used in a pharmaceutical composition also containing
CC	XX	an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC	XX	recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC	XX	does not undergo significant genetic deletion, thereby providing an
CC	XX	unexpectedly stable and immunogenic pox virus. They can be used to
CC	XX	prevent or treat tumours expressing MUC1 tumour-associated antigens

```
XX
SQ Sequence 60 BP; 13 A; 27 C; 17 G; 3 T; 0 U; 0 Other;
Query Match
Best Local Similarity 100.0%; Score 60; DB 2; Length 60;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GGCAGCACCAGCCGCCGACACAGGGGTACAAAGCGCGCCAGACACTGACCTGGCGCA 60
    |||||
    1 GGCAGCACCAGCCGCCGACACAGGGGTACAAAGCGCGCCAGACACTGACCTGGCGCA 60
Db 1 GGCAGCACCAGCCGCCGACACAGGGGTACAAAGCGCGCCAGACACTGACCTGGCGCA 60

RESULT 2
AAV48322
ID AAV48322 standard; DNA; 60 BP.
XX
AC AAV48322;
XX
DT 20-NOV-1998 (first entry)
XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit R7.
XX
KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KM tumour; tumour-associated antigen.
XX
OS Homo sapiens.
XX
PN WO9837095-A2.
XX
PD 27-AUG-1998.
XX
PF 24-FEB-1998; 98WO-US003693.
XX
PR 24-FEB-1997; 97US-0038253P.
XX
PA (THER-) THERION BIOLOGICS CORP.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PA (DAND ) DANA FARBER CANCER INST INC.
XX
PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX
DR WPI; 1998-467492/40.
XX
PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS Disclosure; Page 11; 42pp; English.
XX
CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ Sequence 60 BP; 14 A; 24 C; 16 G; 6 T; 0 U; 0 Other;
Query Match
Best Local Similarity 86.7%; Score 52; DB 2; Length 60;
Matches 55; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
OY 1 GGCAGCACCAGCCGCCGACACAGGGGTACAAAGCGCGCCAGACACTGACCTGGCGCA 60
    |||||
    1 GGCAGCACCAGCCGCCGACACAGGGGTACAAAGCGCGCCAGACACTGACCTGGCGCA 60
Db 1 GGCAGCACCAGCCGCCGACACAGGGGTACAAAGCGCGCCAGACACTGACCTGGCGCA 60

RESULT 3
AAV48325
ID AAV48325 standard; DNA; 60 BP.
XX
AC AAV48325;
```

```
XX
DT 20-NOV-1998 (first entry)
XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit R10.
XX
KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KM tumour; tumour-associated antigen.
XX
OS Homo sapiens.
XX
PN WO9837095-A2.
XX
PD 27-AUG-1998.
XX
PF 24-FEB-1998; 98WO-US003693.
XX
PR 24-FEB-1997; 97US-0038253P.
XX
PA (THER-) THERION BIOLOGICS CORP.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PA (DAND ) DANA FARBER CANCER INST INC.
XX
PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX
DR WPI; 1998-467492/40.
XX
PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS Disclosure; Page 11; 42pp; English.
XX
CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ Sequence 60 BP; 12 A; 22 C; 19 G; 7 T; 0 U; 0 Other;
Query Match
Best Local Similarity 61.0%; Score 36.6; DB 2; Length 60;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
OY 1 GGCAGCACCAGCCGCCGACACAGGGGTACAAAGCGCGCCAGACACTGACCTGGCGCC 59
    |||||
    1 GGCAGCACCAGCCGCCGACACAGGGGTACAAAGCGCGCCAGACACTGACCTGGCGCC 59
Db 1 GGCAGCACCAGCCGCCGACACAGGGGTACAAAGCGCGCCAGACACTGACCTGGCGCC 59

RESULT 4
AAV48317
ID AAV48317 standard; DNA; 60 BP.
XX
AC AAV48317;
XX
DT 20-NOV-1998 (first entry)
XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit R2.
XX
KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KM tumour; tumour-associated antigen.
XX
OS Homo sapiens.
XX
PN WO9837095-A2.
XX
PD 27-AUG-1998.
XX
PF 24-FEB-1998; 98WO-US003693.
XX
PR 24-FEB-1997; 97US-0038253P.
```

CC	an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC	recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC	does not undergo significant genetic deletion, thereby providing an
CC	unexpectedly stable and immunogenic pox virus. They can be used to
CC	prevent or treat tumours expressing MUC1 tumour-associated antigens
XX	
SQ	Sequence 60 BP; 7 A; 29 C; 19 G; 5 T; 0 U; 0 Other;
QY	
	Query Match 53.0%; Score 31.8; DB 2; Length 60;
	Best Local Similarity 71.2%; Pred. No. 0.7;
Db	Matches 42; Conservative 0; Mismatches 17; Indels 0; Gaps 0
	1 GGCAGCACCGCACCGCCCGGTACACAAGCGCCGACGACTGCAGCTGCGCC 59
	1 GGATCCACCGCGCGCCCTGCGCACGAGTGACGTGCGCGCCGACACGCGCCGCTCC 59
RESULT 6	
ID	AAV48324 standard; DNA; 60 BP.
XX	
XX	AAV48324;
XX	
DT	20-NOV-1998 (first entry)
DE	Nucleotide sequence encoding MUC1 tandem repeat unit R9.
XX	
XX	ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW	tumour; tumour-associated antigen.
XX	
OS	Homo sapiens.
XX	
XX	WO9837095-A2.
PN	
XX	27-AUG-1998.
PD	
XX	
PF	24-FEB-1998; 98WO-US003693.
XX	
PR	24-FEB-1997; 97US-0038253P.
XX	
PA	(THER-) THERION BIOLOGICS CORP.
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.
PA	(DAND) DANA FARBER CANCER INST INC.
XX	
PI	Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX	
DR	WPI; 1998-467492/40.
XX	
PT	New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT	immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX	
PS	Disclosure; Page 11; 42pp; English.
XX	
CC	The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC	immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC	(RPV). The RPV was used in a pharmaceutical composition also containing
CC	an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC	recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC	does not undergo significant genetic deletion, thereby providing an
CC	unexpectedly stable and immunogenic pox virus. They can be used to
CC	prevent or treat tumours expressing MUC1 tumour-associated antigens
XX	
SQ	Sequence 60 BP; 12 A; 26 C; 14 G; 8 T; 0 U; 0 Other;
QY	
	Query Match 51.0%; Score 30.6; DB 2; Length 60;
	Best Local Similarity 73.6%; Pred. No. 1.7;
Db	Matches 39; Conservative 0; Mismatches 14; Indels 0; Gaps 0
	7 ACCGACCGCCCGCACAGGGGTACACAAGCGCCGACGACTGCAGCTGCGCC 59
	7 ACGGACCTCCAGCACAGGAGTCAAGTGTGACACCGGACACCGCTCCAGCTCC 59


```
RESULT 7
AAV48316
ID AAV48316 standard; cDNA; 60 BP.
XX
AC AAV48316;
XX
DT 20-NOV-1998 (first entry)
XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit.
XX
KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..60
FT /*tag= a
FT /product= "MUC1 tandem repeat unit"
XX
XX W09837095-A2.
XX
XX PD 27-AUG-1998.
XX
XX PF 24-FEB-1998; 98WO-US003693.
XX
XX PR 24-FEB-1997; 97US-0038253P.
XX
XX PA (THER-) THERION BIOLOGICS CORP.
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PA (DAND ) DANA FARBER CANCER INST INC.
XX
XX PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX
XX DR WPI; 1998-467492/40.
XX DR P-PSDB; AAW77229.
XX
XX PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
XX PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
XX PS Example 1; Page 20; 42pp; English.
XX
XX CC The MUC1 tandem repeat unit was used to create an immunogenic mini-MUC1
XX CC fragment for inclusion in a recombinant pox virus (RPV). The RPV was used
XX CC in a pharmaceutical composition also containing an immunomodulator to
XX CC generate MUC1 specific cytotoxic T-lymphocytes. The recombinant pox virus
XX CC therefore encodes an immunogenic MUC1 fragment that does not undergo
XX CC significant genetic deletion, thereby providing an unexpectedly stable
XX CC and immunogenic pox virus. They can be used to prevent or treat tumours
XX CC expressing MUC1 tumour-associated antigens
XX
SQ Sequence 60 BP; 7 A; 33 C; 16 G; 4 T; 0 U; 0 Other;

Query Match 50.3%; Score 30.2; DB 2; Length 60;
Best Local Similarity 69.5%; Pred. No. 2.2;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

OY 1 GGCAGCACCACCGCCGCGGTCACACAGCGCGCCAGACACTCGACCTGCGCC 59
DB 1 GGCTCCACCGCCCGCCCGACCGCGGTGTCACTCGCGCCCGGACACCAAGCGCGCC 59

RESULT 8
AAV48326
ID AAV48326 standard; cDNA; 78 BP.
XX
AC AAV48326;
XX
DT 20-NOV-1998 (first entry)
XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit a.
XX
KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
```

```
KW tumour; tumour-associated antigen.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX FT CDS 1..78
XX FT /*tag= a
XX FT /product= "MUC1 tandem repeat unit"
XX
XX W09837095-A2.
XX
XX PD 27-AUG-1998.
XX
XX PF 24-FEB-1998; 98WO-US003693.
XX
XX PR 24-FEB-1997; 97US-0038253P.
XX
XX PA (THER-) THERION BIOLOGICS CORP.
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PA (DAND ) DANA FARBER CANCER INST INC.
XX
XX PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX
XX DR WPI; 1998-467492/40.
XX DR P-PSDB; AAW77230.
XX
XX PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
XX PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
XX PS Example 1; Page 20; 42pp; English.
XX
XX CC The MUC1 tandem repeat units AAV48326-V48328 were used to create an
XX CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
XX CC (RPV). The RPV was used in a pharmaceutical composition also containing
XX CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
XX CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
XX CC does not undergo significant genetic deletion, thereby providing an
XX CC unexpectedly stable and immunogenic pox virus. They can be used to
XX CC prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ Sequence 78 BP; 9 A; 42 C; 22 G; 5 T; 0 U; 0 Other;

Query Match 50.3%; Score 30.2; DB 2; Length 78;
Best Local Similarity 69.5%; Pred. No. 2.3;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

OY 1 GGCAGCACCACCGCCGCGGTCACACAGCGCGCCAGACACTCGACCTGCGCC 59
DB 1 GGCTCCACCGCCCGCCCGACCGCGGTGTCACTCGCGCCCGGACACCAAGCGCGCC 59

RESULT 9
AAN90579/C
ID AAN90579 standard; cDNA; 309 BP.
XX
AC AAN90579;
XX
XX DT 27-AUG-2003 (revised)
XX DT 25-MAR-2003 (revised)
XX DT 04-DEC-1989 (first entry)
XX
XX DE PDF9.3 cDNA insert.
XX
XX KW PDF9.3; human DF3 breast carcinoma-associated antigen epitope.
XX
XX OS Human MCF-7 breast carcinoma cells; 'ATCC HTB22'.
XX
XX PN W08907107-A.
XX
XX PD 10-AUG-1989.
XX
XX PF 29-JAN-1988; 88US-00149831.
XX
```

PR 29-JAN-1988; 88US-00149831.
XX
PA (DANA-) DANA-FARBBER CANCER.
XX
PI Kufe DW;
XX
DR WPI; 1989-248989/34.
DR P-PSDB; AAP91045, AAP91053, AAP91054, AAP90146.
XX
PT Recombinant polypeptide(s) - contains DF 3 breast carcinoma antigen
PT epitope and useful as assay reagents, and encoding DNA sequences.
XX
PS Claim 1; Fig 4; 31pp; English.
XX
CC The sequence encodes a carbohydrate-free polypeptide contg. DF3 breast
CC carcinoma antigen epitope. Useful as a competitive binding assay reagent
CC and improves diagnosis. The CDNA consists of nearly identical 60 BP
CC tandem repeats which are 85% GC-rich. See also AAP91045, AAP91046,
CC AAP91053, AAP91054. (Updated on 25-MAR-2003 to correct PA field.)
CC (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 309 BP; 21 A; 92 C; 160 G; 36 T; 0 U; 0 Other;

Query Match 50.3%; Score 30.2; DB 1; Length 309;
Best Local Similarity 69.5%; Pred. No. 2.4;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGCAGACCGCACCGCCCGCACACGGGTCACAAGCGCCGACACTCGACCTGCGCC 59
Db 266 GGCTCCACCGCGCCCGCCCGACCGGTGTCACTCGGCGCCCGGACACCAAGCGCGCCC 208

RESULT 10
AD157712
ID AD157712 standard; cDNA; 1194 BP.
XX
AC AD157712;
XX
DT 22-APR-2004 (first entry)
XX
DE Human breast specific nucleic acid (BSNA) #83.
XX
KW Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
KW breast cancer; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2003106648-A2.
XX
PD 24-DEC-2003.
XX
PF 16-JUN-2003; 2003WO-US018934.
XX
PR 14-JUN-2002; 2002US-0389327P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
XX WPI; 2004-082185/08.
DR P-PSDB; AD157782.
XX
XX
PT Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.
XX
PS Claim 1; SEQ ID NO 83; 370pp; English.
XX
CC The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the

CC hybridisation. The nucleic acids are useful for determining the presence
CC of a BSP in a sample which involves contacting the sample with suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein in the sample
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.
XX
SQ Sequence 1194 BP; 243 A; 438 C; 275 G; 238 T; 0 U; 0 Other;

Query Match 50.3%; Score 30.2; DB 12; Length 1194;
Best Local Similarity 69.5%; Pred. No. 2.6;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGCAGACCGCACCGCCCGCACACGGGTCACAAGCGCCGACACTCGACCTGCGCC 59
Db 596 GGCTCCACCGCGCCCGCCCGACCGGTGTCACTCGGCGCCCGGACACCAAGCGCGCCC 654

RESULT 11
AD157693
ID AD157693 standard; cDNA; 1378 BP.
XX
AC AD157693;
XX
DT 22-APR-2004 (first entry)
XX
DE Human breast specific nucleic acid (BSNA) #64.
XX
KW Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
KW breast cancer; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2003106648-A2.
XX
PD 24-DEC-2003.
XX
PF 16-JUN-2003; 2003WO-US018934.
XX
PR 14-JUN-2002; 2002US-0389327P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
XX WPI; 2004-082185/08.
DR P-PSDB; AD157765.
XX
XX
PT Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.
XX
PS Claim 1; SEQ ID NO 64; 370pp; English.
XX
CC The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the
CC hybridisation. The nucleic acids are useful for determining the presence
CC of a BSP in a sample which involves contacting the sample with suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in

CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein and
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.
XX

SQ Sequence 1378 BP; 295 A; 489 C; 316 G; 278 T; 0 U; 0 Other;

Query Match 50.3%; Score 30.2; DB 12; Length 1378;
Best Local Similarity 69.5%; Pred. No. 2.6;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGCGACACCGCACCGCCCGCACACGGGTACACAGCGCGCCAGACACTCGACTGCGCC 59
DB 596 GGCTCCACCGCCCCCGACGCCAGGTGTCACTCGGCCCGGACACACCGCGCGCCCC 654

RESULT 12

ADO23180/c
ID ADO23180 standard; RNA; 1424 BP.

AC ADO23180;

DT 12-AUG-2004 (first entry)

DE Antisense human MUC1 mucin glycoprotein RNA (coding sequence) SeqID 75.

KW human; MUC1; mucin glycoprotein; cancer; chemotherapeutic; MUC1/ECF; ss;
XX antisense.

OS Homo sapiens.

PN WO2004044160-A2.

PD 27-MAY-2004.

PF 12-NOV-2003; 2003WO-US035848.

PR 13-NOV-2002; 2002US-00293391.

PR 29-MAY-2003; 2003US-00447839.

PA (DAND) DANA FARBER CANCER INST INC.

PA (ILEX-) ILEX PROD INC.

PI Kufe DW, Kharbanda S, Weitman SD;

WPI; 2004-420304/39.

PT Double-stranded RNA complex useful for inhibiting proliferation of cancer
PT cell expressing MUC1 mucin glycoprotein, comprises first and second RNA
PT sequences.

PS Disclosure; SEQ ID NO 75; 112pp; English.

CC This invention relates to novel modulators of the human MUC1 mucin
CC glycoprotein for use in cancer therapeutics, where MUC1 is a protein that
CC acts to inhibit the apoptotic response to genotoxic stress caused by
CC chemotherapeutic agents. In particular, it refers to modulators of the
CC MUC1 extracellular domain (MUC1/ECF). The method refers to using double-
CC stranded RNA complexes as MUC1 interference RNA compositions such that
CC MUC1 expression is inhibited, which in turn inhibits cancer cell
CC proliferation. The present invention describes screening assays to
CC identify compounds that inhibit the binding of various MUC1 ligands such
CC as neuiregulin 2, as well as agonists, antagonists and antibodies thereof.
CC Furthermore, it provides MUC1 antisense, small interfering RNA (siRNA)
CC and small molecules in combination with chemotherapeutic agents that are

CC useful in the field of cancer therapy. This polynucleotide sequence is
CC the antisense human MUC1 RNA of the invention.
XX

SQ Sequence 1424 BP; 309 A; 325 C; 495 G; 0 T; 295 U; 0 Other;

Query Match 50.3%; Score 30.2; DB 12; Length 1424;
Best Local Similarity 69.5%; Pred. No. 2.6;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGCGACACCGCACCGCCCGCACACGGGTACACAGCGCGCCAGACACTCGACTGCGCC 59
DB 1041 GGCTCCACCGCCCCCGACGCCAGGTGTCACTCGGCCCGGACACACCGCGCGCCCC 983

RESULT 13

ABL60159
ID ABL60159 standard; cDNA; 1428 BP.

AC ABL60159;

DT 22-JUL-2002 (first entry)

DE Human MUC1 encoding cDNA SEQ ID NO 2.

KW Human; mucin 1; MUC1; transmembrane protein; SNP; cancer; cytostatic;
KW single nucleotide polymorphism; haplotyping; genotyping; drug;
KW antiinflammatory; gene; ss.

OS Homo sapiens.

Key Location/Qualifiers
1. 1428
/*tag= a
/product= "MUC1"
replace(1009,A)
/*tag= b
/standard name= "Single nucleotide polymorphism"
/note= "SNP allelic variation results in Val substituted
by Met at position 337 of the MUC1 protein (ABB77476) "

FT variation

FT /standard name= "Single nucleotide polymorphism"
FT /note= "SNP allelic variation results in Val substituted
FT by Met at position 337 of the MUC1 protein (ABB77476) "

PN WO200226765-A2.

PD 04-APR-2002.

PF 25-SEP-2001; 2001WO-US030151.

PR 28-SEP-2000; 2000US-0236113P.

PA (GENA-) GENAISSANCE PHARM INC.

PI Chew A, Koshy B;

WPI; 2002-405042/43.

P-PSDB; ABB77476.

PT New genetic variants of mucin 1, Transmembrane gene, useful in studying
PT expression and function of protein encoded by the gene and for screening
PT drugs to treat diseases e.g. cancer.

PS Claim 23; Fig 2; 75pp; English.

CC The invention relates to a polynucleotide (ABL60158, ABL60159) encoding
CC mucin 1/MUC1 (ABB77476), Transmembrane isogene. The invention describes
CC novel genetic variants of the MUC1 gene. The invention is useful for
CC haplotyping/genotyping the MUC1 gene in an individual and identifying an
CC association between a trait and at least one of the haplotypes or
CC haplotype pairs of MUC1 gene. MUC1 is useful for studying the expression
CC and function of MUC1 and expressing MUC1 protein for use in screening for
CC candidate drugs to treat diseases related to MUC1 activity and in
CC studying the effect of the variation on the biological activity of MUC1
CC as well as on the binding affinity of candidate drugs targeting MUC1 for
CC the treatment of e.g. cancer. MUC1 is further used by the pharmaceutical
CC research scientist to validate MUC1 as a candidate target for and in

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 22:37:00 : Search time 183.4 Seconds
(without alignments)
1879.789 Million cell updates/sec

Title: US-10-057-136-7

Perfect score: 60

Sequence: 1 ggctcgacggccccccctgc.....cggataccagacggccccct 60

Scoring table:

IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*

10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*

11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*

13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*

14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*

15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*

16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*

17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*

18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*

19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*

20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	ID	Description
1	60	100.0	60	US-10-057-136-7	Sequence 7, Appli
2	60	100.0	2297	US-10-406-317-41	Sequence 41, Appl
3	43	71.7	60	US-10-057-136-2	Sequence 2, Appli
4	43	71.7	1424	US-10-447-839A-75	Sequence 75, Appl
5	43	71.7	1428	US-10-447-839A-20	Sequence 20, Appl
6	43	71.7	1527	US-10-057-136-19	Sequence 19, Appl
7	43	71.7	1799	US-10-447-839A-19	Sequence 19, Appl
8	43	71.7	1804	US-09-964-824A-573	Sequence 573, Appl
9	43	71.7	1804	US-10-029-517-17	Sequence 17, Appl
10	43	71.7	1804	US-10-717-597-30	Sequence 30, Appl
11	43	71.7	1804	US-10-775-920-84	Sequence 84, Appl
12	43	71.7	1823	US-10-101-510-339	Sequence 339, App

13	43	71.7	4139	9	US-09-964-824A-105	Sequence 105, App
14	43	71.7	4139	9	US-09-964-824A-578	Sequence 578, App
15	43	71.7	4139	9	US-09-864-864-334	Sequence 334, App
16	43	71.7	4139	9	US-09-880-107-2121	Sequence 2121, Ap
17	43	71.7	4139	11	US-09-968-007A-751	Sequence 751, App
18	43	71.7	4139	14	US-10-171-311-157	Sequence 157, App
19	43	71.7	4139	15	US-10-177-293-310	Sequence 310, App
20	43	71.7	4139	16	US-10-440-464-155	Sequence 155, App
21	43	71.7	4139	17	US-10-734-564-53	Sequence 53, Appl
22	43	71.7	4139	17	US-10-775-920-80	Sequence 80, Appl
23	43	71.7	4139	17	US-10-775-920-85	Sequence 85, Appl
24	43	71.7	8186	16	US-10-447-839A-18	Sequence 18, Appl
25	43	71.7	8186	14	US-10-247-703-23	Sequence 23, Appl
26	43	71.7	8186	15	US-10-029-517-19	Sequence 19, Appl
27	41.4	69.0	572	15	US-10-029-517-18	Sequence 18, Appl
28	41.4	69.0	572	17	US-10-775-920-93	Sequence 93, Appl
29	41.4	69.0	1721	9	US-09-864-864-280	Sequence 280, App
30	41.4	69.0	1721	9	US-09-967-768A-224	Sequence 224, App
31	41.4	69.0	1721	14	US-10-247-703-21	Sequence 21, Appl
32	41.4	69.0	1721	14	US-10-097-340-211	Sequence 211, Appl
33	41.4	69.0	1721	14	US-10-171-311-155	Sequence 155, App
34	41.4	69.0	1721	15	US-10-007-926A-58	Sequence 58, Appl
35	41.4	69.0	1721	15	US-10-029-517-3	Sequence 3, Appl1
36	41.4	69.0	1721	15	US-10-172-118-775	Sequence 775, App
37	41.4	69.0	1721	16	US-10-342-887-775	Sequence 775, App
38	41.4	69.0	1721	17	US-10-775-920-88	Sequence 88, Appl
39	41.4	69.0	2026	14	US-10-198-846-12589	Sequence 12589, A
40	41.4	69.0	2238	17	US-10-775-920-87	Sequence 87, Appl
41	41.4	69.0	2678	15	US-10-252-157-103	Sequence 103, App
42	41	68.3	518	14	US-10-247-703-38	Sequence 38, Appl
43	41	68.3	518	15	US-10-029-517-101	Sequence 101, App
44	39.8	66.3	78	14	US-10-057-136-13	Sequence 13, Appl
45	39.8	66.3	3343	14	US-10-247-703-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1

US-10-057-136-7

: Sequence 7, Application US/10057136

: Publication No. US20030021770A1

: GENERAL INFORMATION:

: APPLICANT: SCHLOM, JEFFREY

: APPLICANT: KANTOR, JUDITH

: APPLICANT: KUFFE, DONALD

: APPLICANT: PANICALI, DENNIS

: APPLICANT: GRITZ, LINDA

: TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1

: FILE REFERENCE: 700953/47113C

: CURRENT APPLICATION NUMBER: US/10/057, 136

: PRIOR FILING DATE: 2002-01-25

: PRIOR APPLICATION NUMBER: 09/366, 670

: PRIOR FILING DATE: 1999-08-03

: PRIOR APPLICATION NUMBER: PCT/US98/03693

: PRIOR FILING DATE: 1998-02-24

: PRIOR APPLICATION NUMBER: 60/038, 253

: PRIOR FILING DATE: 1997-02-24

: NUMBER OF SEQ ID NOS: 20

: SOFTWARE: PatentIn Ver. 2.1

: SEQ ID NO 7

: LENGTH: 60

: TYPE: DNA

: ORGANISM: Homo sapiens

US-10-057-136-7

Query Match 100.0%; Score 60; DB 14; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.4e-12;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGTTGACGGCCCCCTGCTACAGGTGTAAACATCGCCCGGATACAGACCGCCCT 60
|||||

Db 1 GGTTCGACGGCCCCCTGCTCAGGTGTAAACATCCGCCCCGGATACAGACCGGCCCT 60

RESULT 2

US-10-406-317-41
; Sequence 41, Application US/10406317
; Publication No. US20040019195A1
; GENERAL INFORMATION:
; APPLICANT: Schlom, Jeffrey;
; APPLICANT: Hodge, James;
; APPLICANT: Panicali, Dennis
; TITLE OF INVENTION: A recombinant vector expressing multiple constitutulatory
; FILE REFERENCE: 38163-0189
; CURRENT APPLICATION NUMBER: US/10/406,317
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US/09/856,988
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: PCT/US99/26866
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/111,582
; PRIOR FILING DATE: 1998-12-09
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: VECTOR
US-10-406-317-41

Query Match 100.0%; Score 60; DB 16; Length 2297;
Best Local Similarity 100.0%; Pred. No. 1.2e-12;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCGACGGCCCCCTGCTCAGGTGTAAACATCCGCCCCGGATACAGACCGGCCCT 60
Db 586 GGTTCGACGGCCCCCTGCTCAGGTGTAAACATCCGCCCCGGATACAGACCGGCCCT 645

RESULT 3

US-10-057-136-2
; Sequence 2, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-2

Query Match 71.7%; Score 43; DB 14; Length 60;
Best Local Similarity 83.1%; Pred. No. 2.6e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GGTTCGACGGCCCCCTGCTCAGGTGTAAACATCCGCCCCGGATACAGACCGGCCCT 59
Db 1 GGTTCGACGGCCCCCTGCTCAGGTGTAAACATCCGCCCCGGATACAGACCGGCCCT 59

RESULT 4

US-10-447-839A-75/c
; Sequence 75, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald W.
; APPLICANT: Kharbanda, Sunder
; APPLICANT: Weitman, Steven D.
; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT APPLICATION NUMBER: US/10/447,839A
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 75
; LENGTH: 1424
; TYPE: RNA
; ORGANISM: ARTIFICIAL
; FEATURE:
; OTHER INFORMATION: Synthesized Sequence
US-10-447-839A-75

Query Match 71.7%; Score 43; DB 16; Length 1424;
Best Local Similarity 83.1%; Pred. No. 2.3e-06;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GGTTCGACGGCCCCCTGCTCAGGTGTAAACATCCGCCCCGGATACAGACCGGCCCT 59
Db 1041 GGCTCCACCGCCCCCTGCTCAGGTGTAAACATCCGCCCCGGATACAGACCGGCCCT 983

RESULT 5

US-10-447-839A-20
; Sequence 20, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald W.
; APPLICANT: Kharbanda, Sunder
; APPLICANT: Weitman, Steven D.
; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT APPLICATION NUMBER: US/10/447,839A
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 1428
; TYPE: RNA
; ORGANISM: RNA
US-10-447-839A-20

Query Match 71.7%; Score 43; DB 16; Length 1428;
Best Local Similarity 76.3%; Pred. No. 2.3e-06;
Matches 45; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

Dy 1 GGTTCGACGGCCCCCTGCTCACGGTGTAAATCCGCCCGGATTACCAGACGGCCCC 59
||:|||||||:|||||:
Ddb 385 GGCUCACCGCCCCCAAGCCACGGUGUACAUCUGGCCCGCAGAACCAAGCGCGCCCC 443

RESULT 6

```

US-10-057-136-19
; Sequence 19, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOW, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; TITLE OF INVENTION: TUMOR-ASSOCIATED ANTIGEN
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 1527
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-19

```

Query Match	71.7%;	Score 43;	DB 14;	Length 1527;
Best Local Similarity	83.1%;	Pred. No. 2.3e-06;		
Matches 49;	Conservative 0;	Mismatches 10;	Indels 0;	Gaps 0;

```
Oy      1 GGTTCGACGGCCCCCTGCTCACGGTGTAAATCCGCCCGGATTACCAGACCGGCCCC 59
        |||||
Db     226 GGCTTCAACGGCCCCCAGACCCAAGTGTACTCGGCCCGGACACCAAGCGCGGCCCC 284
```

RESULT 7

```

US-10-447-839A-19
; Sequence 19, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald W.
; APPLICANT: Kharbanda, Surender
; APPLICANT: Weltman, Steven D.
; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT APPLICATION NUMBER: US/10/447,839A
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 1799
; TYPE: RNA
; ORGANISM: RNA
US-10-447-839A-19

```

Query Match	71.7%;	Score 43;	DB 16;	Length 1799;
Best Local Similarity	76.3%;	Pred. No. 2.2e-06;		
Matches 45;	Conservative 4;	Mismatches 10;	Indels 0;	Gaps 0;

[illegible]

RESULT 8

```

US-09-964-824A-573
; Sequence 573, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horrigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964,824A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 573
; LENGTH: 1804
;
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-824A-573

```

Query Match	71.7%	Score 43;	DB 9;	Length 1804;
Best Local Similarity	83.1%	Pred. No. 2.2e-06;		
Matches 49; Conservative	0;	Mismatches 10;	Indels 0;	Gaps 0;

```
Oy      1 GGTTCAGCGCCCCCTGCTACGGTGTAAATCCGCCCGGATAACAGACCGGCCCC 59  
         ||| |||| |  
Db    457 GGCTCCACCGCCCCCAGGCCCAAGTGTCACCTCGGCCCGGACAACAGGCCGCCCC 515
```

RESULT 9

```

US-10-029-517-17
; Sequence 17, Application US/10029517
; Publication No. US20030148969A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 17
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (73)... (1500)
US-10-029-517-17

```

Query Match	71.7%;	Score 43;	DB 15;	Length 1804;
Best Local Similarity	83.1%;	Pred. No. 2.2e-06;		
Matches 49; Conservative	0;	Mismatches 10;	Indels 0;	Gaps 0;

[illegible]

RESULT 10

US-10-717-597-30
; Sequence 30, Application US/10717597

```
; Publication No. US20040110221A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael E.
; APPLICANT: Twine, Natalie C.
; APPLICANT: Dorneer, Andrew J.
; APPLICANT: Trepicchio, William L.
; APPLICANT: Slonim, Donna K.
; APPLICANT: Stover, Jennifer A.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING RCC AND OTHER SOLID TUMORS
; FILE REFERENCE: AM101080L
; CURRENT APPLICATION NUMBER: US/10/717,597
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US 60/459,782
; PRIOR FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: US 60/427,982
; PRIOR FILING DATE: 2002-11-21
; NUMBER OF SEQ ID NOS: 4904
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-717-597-30
```

```
Query Match          71.7%; Score 43; DB 17; Length 1804;
Best Local Similarity 83.1%; Pred. No. 2.2e-06;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 1 GGTTCGACGGCCCCCTGCTCAGCGGTGTACATCCGCGCGATACAGACGGCCCC 59
    |||||
Db 457 GGCTCCACCGCCCCCGACGCCACCGGTGTCACTCGGCCCGGACACACGCGCGCCCC 515
```

```
RESULT 11
US-10-775-920-84
; Sequence 84, Application US/10775920
; Publication No. US20040175744A1
; GENERAL INFORMATION:
; APPLICANT: Merzen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; FILE REFERENCE: Merzen - 0010B
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 84
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-775-920-84
```

```
Query Match          71.7%; Score 43; DB 17; Length 1804;
Best Local Similarity 83.1%; Pred. No. 2.2e-06;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 1 GGTTCGACGGCCCCCTGCTCAGCGGTGTACATCCGCGCGATACAGACGGCCCC 59
    |||||
Db 457 GGCTCCACCGCCCCCGACGCCACCGGTGTCACTCGGCCCGGACACACGCGCGCCCC 515
```

```
RESULT 12
US-10-101-510-339
; Sequence 339, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
```

```
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 339
; LENGTH: 1823
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-101-510-339
```

```
Query Match          71.7%; Score 43; DB 15; Length 1823;
Best Local Similarity 83.1%; Pred. No. 2.2e-06;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 1 GGTTCGACGGCCCCCTGCTCAGCGGTGTACATCCGCGCGATACAGACGGCCCC 59
    |||||
Db 457 GGCTCCACCGCCCCCGACGCCACCGGTGTCACTCGGCCCGGACACACGCGCGCCCC 515
```

```
RESULT 13
US-09-964-824A-105
; Sequence 105, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horriagan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signat
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964,824A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 105
; LENGTH: 4139
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-964-824A-105
```

```
Query Match          71.7%; Score 43; DB 9; Length 4139;
Best Local Similarity 83.1%; Pred. No. 2.2e-06;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 1 GGTTCGACGGCCCCCTGCTCAGCGGTGTACATCCGCGCGATACAGACGGCCCC 59
    |||||
Db 458 GGCTCCACCGCCCCCGACGCCACCGGTGTCACTCGGCCCGGACACACGCGCGCCCC 516
```

```
RESULT 14
US-09-964-824A-578
; Sequence 578, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horriagan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964,824A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
; PRIOR FILING DATE: 2000-09-28
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: January 15, 2005, 19:50:13 ; Search time 35.2 Seconds
(without alignments)
1211.572 Million cell updates/sec

Title: US-10-057-136-7

Perfect score: 60
Sequence: 1 ggttcgacggccccctgc.....cggataccagacggccct 60

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	43	71.7	1804	4 US-10-029-517-17 Sequence 17, Appl
2	43	71.7	8186	4 US-10-029-517-19 Sequence 19, Appl
3	41.4	69.0	572	4 US-10-029-517-18 Sequence 18, Appl
4	41.4	69.0	1721	4 US-10-029-517-3 Sequence 3, Appl
5	41	68.3	518	4 US-10-029-517-101 Sequence 101, App
6	39.8	66.3	3343	4 US-10-029-517-102 Sequence 102, App
7	38	63.3	60	4 US-09-475-947A-246 Sequence 246, App
8	36.2	60.3	981	4 US-10-029-517-16 Sequence 16, Appl
9	35.6	59.3	6192	2 US-08-479-537A-1 Sequence 1, Appl
10	35.6	59.3	6192	3 US-09-083-116-1 Sequence 1, Appl
11	35.6	59.3	6192	3 US-09-134-916A-1 Sequence 1, Appl
12	35.6	59.3	6449	2 US-08-479-537A-4 Sequence 4, Appl
13	35.6	59.3	6449	3 US-09-083-116-4 Sequence 4, Appl
14	35.6	59.3	6449	3 US-09-134-916A-4 Sequence 4, Appl
15	26.6	44.3	519	4 US-09-646-028-42 Sequence 42, Appl
16	26.6	44.3	534	4 US-09-646-028-46 Sequence 46, Appl
17	25.6	42.7	392	4 US-09-513-999C-2024 Sequence 2024, App
18	24.4	40.7	556	4 US-10-029-517-105 Sequence 105, App
19	24.2	40.3	38653	4 US-09-922-445-1 Sequence 1, Appl
20	23	38.3	3762	4 US-09-435-376-1 Sequence 1, Appl
21	22.8	38.0	1737	2 US-08-750-703-2 Sequence 2, Appl
22	22.8	38.0	8779	2 US-08-750-703-4 Sequence 4, Appl
23	22.8	38.0	77536	4 US-09-410-551B-1 Sequence 1, Appl
24	22.8	38.0	77536	4 US-09-410-551B-1 Sequence 1, Appl
25	22.8	38.0	4403765	3 US-09-103-840A-2 Sequence 2, Appl
26	22.8	38.0	4411529	3 US-09-103-840A-1 Sequence 1, Appl
27	22.6	37.7	3188	3 US-08-943-731-183 Sequence 183, App

C	28	22.6	37.7	20084	3	US-08-943-731-5	Sequence 5, Appl
	29	22.4	37.3	3900	4	US-09-023-655-1420	Sequence 1420, Ap
C	30	22.4	37.3	4403765	3	US-09-103-840A-2	Sequence 2, Appl
C	31	22.4	37.3	4411529	3	US-09-103-840A-1	Sequence 1, Appl
	32	22.2	37.0	786	4	US-09-252-991A-12436	Sequence 12436, A
C	33	22.2	37.0	1206	4	US-09-252-991A-13220	Sequence 13220, A
	34	22.2	37.0	1719	4	US-09-252-991A-12741	Sequence 12741, A
	35	22.2	37.0	1806	4	US-09-865-879-1	Sequence 1, Appl
	36	22.2	37.0	2050	4	US-09-891-053-21	Sequence 21, Appl
	37	22.2	37.0	2689	2	US-08-985-090-1	Sequence 1, Appl
	38	22.2	37.0	2689	3	US-09-165-543-1	Sequence 5, Appl
	39	22.2	37.0	2699	3	US-09-167-354-5	Sequence 5, Appl
	40	22.2	37.0	2699	4	US-09-642-855-5	Sequence 5, Appl
	41	22.2	37.0	2699	4	US-09-642-514-5	Sequence 5, Appl
C	42	22.2	37.0	47981	4	US-09-679-279-1	Sequence 1, Appl
	43	22	36.7	8285	4	US-09-732-025-3	Sequence 3, Appl
	44	22	36.7	11827	4	US-09-739-455-3	Sequence 3, Appl
C	45	21.8	36.3	143	4	US-09-513-999C-10197	Sequence 10197, A

ALIGNMENTS

RESULT 1
US-10-029-517-17
; Sequence 17, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 17
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (73)...(1500)
US-10-029-517-17

Query Match
Best Local Similarity 83.1%; Pred. No. 4.1e-06;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GGTTCGACGGCCCCCTGCTCAGCGTGTACATCCGCCCGGATACCAGACGGCCCC 59
Db 457 GGCTCCACCGCCCCCGCCAGCCAGCGTGTACCTCGGGCCCCGACACAGCGCGCCCC 515

RESULT 2
US-10-029-517-19
; Sequence 19, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 19
; LENGTH: 8186
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 6899
; OTHER INFORMATION: unknown

```
; NAME/KEY: unsure
; LOCATION: 7155
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7184
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7957
; OTHER INFORMATION: unknown
; NAME/KEY: intron
; LOCATION: (2997)...(3498)
; OTHER INFORMATION: intron 1
; NAME/KEY: intron:exon junction
; LOCATION: (3498)...(3499)
; OTHER INFORMATION: intron 1:exon 2
; NAME/KEY: exon
; LOCATION: (3508)...(3599)
; OTHER INFORMATION: exon 2d
; NAME/KEY: exon:intron junction
; LOCATION: (3982)...(3983)
; OTHER INFORMATION: exon 2a:intron 2a
; NAME/KEY: intron:exon junction
; LOCATION: (4205)...(4206)
; OTHER INFORMATION: intron 2c:exon 3c
; NAME/KEY: intron:exon junction
; LOCATION: (4259)...(4260)
; OTHER INFORMATION: intron 2d:exon 3d
; NAME/KEY: exon
; LOCATION: (4260)...(4328)
; OTHER INFORMATION: exon 3d
; NAME/KEY: intron:exon junction
; LOCATION: (4632)...(4633)
; OTHER INFORMATION: intron 3:exon 4
; NAME/KEY: exon
; LOCATION: (4914)...(5035)
; OTHER INFORMATION: exon 5
; NAME/KEY: intron
; LOCATION: (5266)...(6293)
; OTHER INFORMATION: intron 6
US-10-029-517-19
```

Query Match 71.7%; Score 43; DB 4; Length 8186;
Best Local Similarity 83.1%; Pred. No. 4.9e-06;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Oy 1 GGTTCGACGGCCCCCTGCTCAGCGGTGAACATCCGCGCCCGGATACCAAGACCGGCCCC 59
Db 3825 GGCTCCACCGCCCCCGCCAGCCACGGGTGTCACTCGGCCCCCGACACCAAGCGCGCCCC 3883
```

```
RESULT 3
US-10-029-517-18
; Sequence 18, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 18
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (67)...(572)
US-10-029-517-18
```

Query Match 69.0%; Score 41.4; DB 4; Length 572;
Best Local Similarity 81.4%; Pred. No. 1.3e-05;

```
Matches 48; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
Oy 1 GGTTCGACGGCCCCCTGCTCAGCGGTGAACATCCGCGCCCGGATACCAAGACCGGCCCC 59
Db 478 GGCTCCACCGCCCCCAAGCCACGGGTGTCACTCGGCCCCCGACACCAAGCGCGCCCC 536
```

```
RESULT 4
US-10-029-517-3
; Sequence 3, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 3
; LENGTH: 1721
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (58)...(1605)
US-10-029-517-3
```

Query Match 69.0%; Score 41.4; DB 4; Length 1721;
Best Local Similarity 81.4%; Pred. No. 1.5e-05;
Matches 48; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```
Oy 1 GGTTCGACGGCCCCCTGCTCAGCGGTGAACATCCGCGCCCGGATACCAAGACCGGCCCC 59
Db 442 GGCTCCACCGCCCCCGCCAGCCACGGGTGTCACTCGGCCCCCGACACCAAGCGCGCCCC 500
```

```
RESULT 5
US-10-029-517-101
; Sequence 101, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 101
; LENGTH: 518
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-10-029-517-101
```

Query Match 68.3%; Score 41; DB 4; Length 518;
Best Local Similarity 82.5%; Pred. No. 1.8e-05;
Matches 47; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Oy 1 GGTTCGACGGCCCCCTGCTCAGCGGTGAACATCCGCGCCCGGATACCAAGACCGGCCC 57
Db 462 GGCTCCACCGCCCCCGCCAGCCACGGGTGTCACTCGGCCCCCGACACCAAGCGCGCCC 518
```

```
RESULT 6
US-10-029-517-102
; Sequence 102, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
```

```
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 102
; LENGTH: 3343
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-10-029-517-102
```

```
Query Match          66.3%; Score 39.8; DB 4; Length 3343;
Best Local Similarity 79.7%; Pred. No. 5.7e-05;
Matches 47; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
```

```
QY      1 GGTTCGACGGCCCCCTGCTCAGCGGTGAACATCCGCCCGGATACAGACCGGCC 59
          |||||
Db      1668 GGCTCCACCGCCCCCGACGCCAGGGTGTCACTCGGCCCGGAGACCAAGCGCCCC 1726
```

```
RESULT 7
US-09-475-947A-246
; Sequence 246, Application US/09475947A
; Patent No. 6472154
; GENERAL INFORMATION:
; APPLICANT: Garner, Harold R.
; APPLICANT: Wren, Jonathan D.
; APPLICANT: Minna, John D.
; TITLE OF INVENTION: Polymorphic Repeats in Human Genes
; FILE REFERENCE: UTS0667
; CURRENT APPLICATION NUMBER: US/09/475,947A
; CURRENT FILING DATE: 1999-12-31
; NUMBER OF SEQ ID NOS: 346
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 246
; LENGTH: 60
; TYPE: DNA
; ORGANISM: human
US-09-475-947A-246
```

```
Query Match          63.3%; Score 38; DB 4; Length 60;
Best Local Similarity 81.5%; Pred. No. 0.00015;
Matches 44; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

```
QY      1 GGTTCGACGGCCCCCTGCTCAGCGGTGAACATCCGCCCGGATACAGACCG 54
          |||||
Db      7 GGCTCCACCGCCCCCGACGCCAGGGTGTCACTCGGCCCGGAGACCAAGCGCC 60
```

```
RESULT 8
US-10-029-517-16
; Sequence 16, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 16
; LENGTH: 981
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon:exon junction
; LOCATION: (464)...(465)
; OTHER INFORMATION: exon 3b:exon 4
US-10-029-517-16
```

```
Query Match          60.3%; Score 36.2; DB 4; Length 981;
Best Local Similarity 77.2%; Pred. No. 0.00088;
```

```
Matches 44; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY      1 GGTTCGACGGCCCCCTGCTCAGCGGTGAACATCCGCCCGGATACAGACCGGCC 57
          |||||
Db      21 GGCTCCACCGCCCCCGACGCCAGGGTGTCACTCGGCCCGGAGACCAAGCGCGCC 77
```

```
RESULT 9
US-08-479-537A-1
; Sequence 1, Application US/08479537A
; Patent No. 5861381
; GENERAL INFORMATION:
; APPLICANT: CHAMON, Pierre
; APPLICANT: KIENY, Marie-Paule
; APPLICANT: LATHE, Richard
; APPLICANT: HAREUVENI, Mara
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
; TITLE OF INVENTION: TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,537A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 90/13101
; FILING DATE: 23-OCT-1990
; APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR91/00835
; FILING DATE: 23-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/039,320
; FILING DATE: 04-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,576
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teekin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 017753-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6192 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 58..120
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 439..5239
; OTHER INFORMATION: /note="The nucleotides spanning
; OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
; OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
; OTHER INFORMATION: The number of such repeats varies from 1 to 80."
; FEATURE:
; NAME/KEY: mat_peptide
```


LOCATION: 121..6166
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-08-479-537A-1

Query Match
Best Local Similarity 59.3%; Score 35.6; DB 2; Length 6192;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGTTGACGGCCCCCTGCTCAGCGGTGTACATCCGCGCGGATACGACCGGCCCC 59
DB 442 GGCTCCACGGCCCCCANNNGCCCAAGGTGTACCTCGCGCCCGGACNNNAGCGCGNNCC 500

RESULT 10
US-09-083-116-1

; Sequence 1, Application US/09083116
; Patent No. 6203795
; GENERAL INFORMATION:
; APPLICANT: CHAMBER, Pierre
; APPLICANT: KIENY, Marie-Paule
; APPLICANT: LATHE, Richard
; APPLICANT: HAREUVENI, Mara
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
; TITLE OF INVENTION: TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/083,116
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,537
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR91/00835
; FILING DATE: 23-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/039,320
; FILING DATE: 04-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,576
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030

REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6192 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note= "The nucleotides spanning
OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixe
OTHER INFORMATION: The number of such repeats varies from 1 to 80."
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 121..6166
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CC
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-09-083-116-1

Query Match
Best Local Similarity 59.3%; Score 35.6; DB 3; Length 6192;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGTTGACGGCCCCCTGCTCAGCGGTGTACATCCGCGCGGATACGACCGGCCCC 59
DB 442 GGCTCCACGGCCCCCANNNGCCCAAGGTGTACCTCGCGCCCGGACNNNAGCGCGNNCC 500

RESULT 11
US-09-134-916A-1

; Sequence 1, Application US/09134916A
; Patent No. 6328956
; GENERAL INFORMATION:
; APPLICANT: CHAMBER, Pierre
; APPLICANT: KIENY, Marie-Paule
; APPLICANT: LATHE, Richard
; APPLICANT: HAREUVENI, Mara
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
; TITLE OF INVENTION: TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:

```

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/134,916A
  FILING DATE:
  CLASSIFICATION:
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US/08/479,537
    FILING DATE: 07-JUN-1995
    APPLICATION NUMBER: FR 90/13101
    FILING DATE: 23-OCT-1990
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: WO PCT/FR91/00835
    FILING DATE: 23-OCT-1991
    APPLICATION NUMBER: US 08/039,320
    FILING DATE: 04-APR-1993
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/403,576
    FILING DATE: 14-MAR-1995
  ATTORNEY/AGENT INFORMATION:
    NAME: Teekin, Robin L.
    REGISTRATION NUMBER: 35,030
    REFERENCE/DOCKET NUMBER: 017753-025
    TELECOMMUNICATION INFORMATION:
      TELEPHONE: (703) 836-6620
      TELEFAX: (703) 836-2021
    INFORMATION FOR SEQ ID NO: 1:
      SEQUENCE CHARACTERISTICS:
        LENGTH: 6192 base pairs
        TYPE: nucleic acid
        STRANDEDNESS: single
        TOPOLOGY: linear
      MOLECULE TYPE: DNA (genomic)
      FEATURE:
        NAME/KEY: sig_peptide
        LOCATION: 58..120
      FEATURE:
        NAME/KEY: repeat_region
        LOCATION: 439..5239
      OTHER INFORMATION: /note= "The nucleotides spanning
      OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
      OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
      OTHER INFORMATION: The number of such repeats varies from 1 to 80."
      FEATURE:
        NAME/KEY: mat_peptide
        LOCATION: 121..6166
      FEATURE:
        NAME/KEY: repeat_region
        LOCATION: 457
      OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
      OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
      OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
      FEATURE:
        NAME/KEY: repeat_region
        LOCATION: 487
      OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
      OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
      OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
      FEATURE:
        NAME/KEY: repeat_region
        LOCATION: 496
      OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
      OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
      OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-09-134-916A-1

```

```

Query Match      59.3%; Score 35.6; DB 3; Length 6192;
Best Local Similarity 69.5%; Pred. No. 0.0018;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```

```

QY      1 GGTTCACGCGCCCCCTGCTCAGCGTGTACATCCGCCCCGATACACGACCGGCCCC 59
Db      442 GGCTCACCGCGCCCCNNNGCCACCGGTGTACCTCGCCCCGACGACGCGGNNCC 500

RESULT 12
US-08-479-537A-4
; Sequence 4, Application US/08479537A
; Patent No. 5861381
; GENERAL INFORMATION:
; APPLICANT: CHAMON, Pierre
; APPLICANT: KIENY, Marie-Paule
; APPLICANT: LATHE, Richard
; APPLICANT: HAREUVENI, Mara
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
; TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,537A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 90/13101
; FILING DATE: 23-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR91/00835
; FILING DATE: 23-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/039,320
; FILING DATE: 04-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,576
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teekin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 017753-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6449 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 58..120
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 439..5239
; OTHER INFORMATION: /note= "The nucleotides spanning
; OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
; OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
; OTHER INFORMATION: The number of such repeats varies from 1 to 80."
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 121..5661
; FEATURE:

```

NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG, and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG, and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG, and Ala = GCT, GCC, GCA, or GCG."
US-08-479-537A-4

Query Match 59.3%; Score 35.6; DB 2; Length 6449;
Best Local Similarity 69.5%; Pred. No. 0.0018;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

OY 1 GGTTGACGGCCCCCTGCTCAGCGGTGTATCCGCGCCGGATACGACGCGGCC 59
Db 442 GGCTCCACCGCCCCCNNGCCGCCACGCTGTCACTCGGCCCGGACNNNAGCGCCGNNCC 500

RESULT 13

US-09-083-116-4
Sequence 4, Application US/09083116
Patent No. 6203795
GENERAL INFORMATION:
APPLICANT: CHAMBERON, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREUVENI, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/083,116
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,537
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6449 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note= "The nucleotides spanning
OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
OTHER INFORMATION: The number of such repeats varies from 1 to 80."

NAME/KEY: mat_peptide
LOCATION: 121..5661
FEATURE:

NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG, and Ala = GCT, GCC, GCA, or GCG."
FEATURE:

NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG, and Asn = AAT or AAC."
FEATURE:

NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG, and Ala = GCT, GCC, GCA, or GCG."
US-09-083-116-4

Query Match 59.3%; Score 35.6; DB 3; Length 6449;
Best Local Similarity 69.5%; Pred. No. 0.0018;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

OY 1 GGTTGACGGCCCCCTGCTCAGCGGTGTATCCGCGCCGGATACGACGCGGCC 59
Db 442 GGCTCCACCGCCCCCNNGCCGCCACGCTGTCACTCGGCCCGGACNNNAGCGCCGNNCC 500

RESULT 14

US-09-134-916A-4
Sequence 4, Application US/09134916A
Patent No. 6328956
GENERAL INFORMATION:
APPLICANT: CHAMBERON, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREUVENI, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```

: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
:   APPLICATION NUMBER: US/09/134,916A
:   FILING DATE:
:     CLASSIFICATION:
:   PRIOR APPLICATION DATA:
:     APPLICATION NUMBER: US/08/479,537
:     FILING DATE: 07-JUN-1995
:     APPLICATION NUMBER: FR 90/13101
:     FILING DATE: 23-OCT-1990
:     PRIOR APPLICATION DATA:
:       APPLICATION NUMBER: WO PCT/FR91/00835
:       FILING DATE: 23-OCT-1991
:     PRIOR APPLICATION DATA:
:       APPLICATION NUMBER: US 08/039,320
:       FILING DATE: 04-APR-1993
:     PRIOR APPLICATION DATA:
:       APPLICATION NUMBER: US 08/403,576
:       FILING DATE: 14-MAR-1995
:     ATTORNEY/AGENT INFORMATION:
:       NAME: Teskin, Robin L.
:       REGISTRATION NUMBER: 35,030
:       REFERENCE/DOCKET NUMBER: 017753-025
:     TELECOMMUNICATION INFORMATION:
:       TELEPHONE: (703) 836-6620
:       TELEFAX: (703) 836-2021
:     INFORMATION FOR SEQ ID NO: 4:
:       SEQUENCE CHARACTERISTICS:
:         LENGTH: 6449 base pairs
:         TYPE: nucleic acid
:         STRANDEDNESS: single
:         TOPOLOGY: linear
:       MOLECULE TYPE: DNA (genomic)
:     FEATURE:
:       NAME/KEY: sig_peptide
:       LOCATION: 58..120
:     FEATURE:
:       NAME/KEY: repeat_region
:       LOCATION: 439..5239
:       OTHER INFORMATION: /note= "The nucleotides spanning
:         OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
:         OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
:         OTHER INFORMATION: The number of such repeats varies from 1 to 80."
:     FEATURE:
:       NAME/KEY: mat_peptide
:       LOCATION: 121..5661
:     FEATURE:
:       NAME/KEY: repeat_region
:       LOCATION: 457
:       OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
:         OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
:         OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
:     FEATURE:
:       NAME/KEY: repeat_region
:       LOCATION: 487
:       OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
:         OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACP
:         OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
:     FEATURE:
:       NAME/KEY: repeat_region
:       LOCATION: 496
:       OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
:         OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCP
:         OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-09-134-916A-4

```

```

Db          442  GGCTCCACCGCCCCNNNGCCCACGGTGTCTACCTCGGCCCCGGACNNNAGGCCGNNCC 500

RESULT 15
US-09-646-028-42
; Sequence 42, Application US/09646028
; Patent No. 6562347
; GENERAL INFORMATION:
; APPLICANT: Kwak, Larry
; APPLICANT: Bliragyn, Arya
; TITLE OF INVENTION: METHODS AND COMPOSITIONS OF
; TITLE OF INVENTION: CHEMOKINE-TUMOR ANTIGEN FUSION PROTEINS AS CANCER VACCINES
; FILE REFERENCE: 14014.0316/P
; CURRENT APPLICATION NUMBER: US/09/646, 028
; CURRENT FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: 60/077, 745
; PRIOR FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 42
; LENGTH: 519
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of artificial sequence:/note=synthetic construct
US-09-646-028-42

```

Query Match	44.3%	Score 26.6	DB 4	Length 519
Best Local Similarity	78.0%	Pred. No. 1.7		
Matches 32	Conservative 0	Mismatches 9	Indels 0	Gaps 0
Qy	1	GGTTGACGGCCCCCTGCTACGGGTGAACATCGCCCC	41	
Db	331	GGTTCTACTGCTCCGCCCGGCACACGGGTGTAACTTCTGCGCG	371	

Search completed: January 16, 2005, 03:01:37
Job time : 44.2 secs

Query Match	59.3%	Score 35.6	DB 3	Length 6449
Best Local Similarity	69.5%	Pred. No. 0.0018		
Matches 41	Conservative 0	Mismatches 18	Indels 0	Gaps 0

QY 1 GGTTCGACGGCCCCCTGCTCAGGTGTACATCCGCCCCGGATACCAGACCGCCCC 59

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 19:43:55 : Search time 1550.6 Seconds
(without alignments)
1410.024 Million cell updates/sec

Title: US-10-057-136-7
Perfect score: 60
Sequence: 1 ggttcgacgccccccctgc.....cggataccagacgccccct 60

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	43	71.7	330	1	AI925867 wo20d04.x
2	43	71.7	604	4	BM791359 K-EST0071
3	43	71.7	877	5	BU542454 AGENCOURT
4	43	71.7	959	6	CA489836 AGENCOURT
5	43	71.7	1113	5	BU148487 AGENCOURT
6	43	71.7	1130	5	BU542996 AGENCOURT
7	43	71.7	1234	5	BQ936898 AGENCOURT
8	43	71.7	1268	5	BQ943554 AGENCOURT
9	43	71.7	1343	5	BQ920055 AGENCOURT
10	43	71.7	1349	5	BU152566 AGENCOURT
11	43	71.7	1420	5	BU542790 AGENCOURT
12	43	71.7	1531	5	BU543309 AGENCOURT
13	41.4	69.0	1262	5	BQ935496 AGENCOURT
14	36.8	61.3	1334	5	BQ943809 AGENCOURT
15	36.2	60.3	981	4	BG774910 602649832
16	36.2	60.3	1536	5	BQ923149 AGENCOURT
17	35.2	58.7	166	7	T27692 EST12384 Hu
18	35.2	58.7	475	6	CB120860 K-EST0168
19	35.2	58.7	669	6	CB122585 K-EST0170
20	35	58.3	1678	4	BG775565 602650481
21	33.8	56.3	619	4	BI260921 602970962
22	33.4	55.7	472	4	BM759495 K-EST0039
c 23	28.6	47.7	878	2	BF786059 602110629
24	28	46.7	635	6	CA924184 MTU7CL.P2

25	28	46.7	635	6	CA924567	MTU7CL.P7
c 26	27.2	45.3	495	2	BF825165	BF825165 RC3-HN002
c 27	27.2	45.3	504	1	AJ280697	AJ280697 4A3A-AAS-
c 28	27	45.0	811	4	BI956680	HVSMEN000
c 29	27	45.0	886	9	CL653412	PR10118b
c 30	26.6	44.3	308	7	R48730	R48730 yj68a12.r1
c 31	26.6	44.3	478	4	BM427484	pgf2n.pk0
c 32	26.6	44.3	832	4	BI518647	BI518647 603062003
c 33	26.6	44.3	1006	5	BX337143	BX337143 BX337143
34	26.4	44.0	750	4	BI523776	BI523776 603052074
35	26.4	44.0	984	4	BI523854	BI523854 603051982
c 36	26.2	43.7	1868	4	BG336063	BG336063 602404834
c 37	26	43.3	548	2	BE706360	BE706360 RC1-HT025
c 38	26	43.3	607	9	CE187066	CE187066 t1gr-gss-
39	26	43.3	682	9	AG082656	AG082656 Pan t1ogr1
40	26	43.3	1119	4	BM911965	BM911965 AGENCOURT
41	25.6	42.7	346	5	BY098559	BY098559 BY098559
42	25.6	42.7	435	8	BH253383	BH253383 SALK 0148
43	25.6	42.7	467	9	CC719739	CC719739 OGRAD34TH
44	25.6	42.7	549	7	CF482287	CF482287 POL1_6 C0
45	25.6	42.7	653	9	CC623364	CC623364 OGUIH90TH

ALIGNMENTS

RESULT 1
AI925867/c 330 bp mRNA linear EST 08-MAR-2000
LOCUS wo20d04.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2455879 3'
DEFINITION similar to contains element MSRI MSRI repetitive element ;, mRNA
sequence.

ACCESSION AI925867
VERSION AI925867
KEYWORDS AI925867.1 GI:5661831
SOURCE EST.

ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 330)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index

JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-rc@mail.nih.gov

Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www-bio.lnl.gov/bbrp/image/image.html
Insert Length: 578 Std Error: 0.00
Seg primer: -40UP from Gibco
High quality sequence stop: 305.

FEATURES
source location/Qualifiers
1..330
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2455879"
/tissue_type="adenocarcinoma"
/lab_host="DH108"
/clone_lib="NCI_CGAP_pan1"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.72 kb. Life Technologies catalog #:
11548-013"

ORIGIN

Query Match 71.7%; Score 43; DB 1; Length 330;
Best Local Similarity 83.1%; Pred. No. 0.00058;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GGTTCGACGGCCCCCTGCTCAGCGGTGTATACATCCGCCCGGATACCAACCGGCCCC 59
|||
Db 316 GGCTCCACCGCCCCCGACGCCACGCGTGTACCTCGGCCCGGACACACGCGCGCCCC 258
|||

RESULT 2 604 bp mRNA linear EST 05-MAR-2002
LOCUS BM791359
DEFINITION K-EST0071342 S21SNUS20 Homo sapiens cDNA clone S21SNUS20-14-A06 5',
mRNA sequence.
BM791359
BM791359.1 GI:19139591

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.krribb.re.kr
Plate: 14 row: A column: 06
High quality sequence stop: 604.
Location/Qualifiers

FEATURES
source
1.604
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S21SNUS20-14-A06"
/sex="F"
/tissue_type="Stomach"
/cell_line="Floating aggregates"
/lab_host="Top10F"
/clone_lib="S21SNUS20"

ORIGIN
Query Match 71.7%; Score 43; DB 4; Length 604;
Best Local Similarity 83.1%; Pred. No. 0.00059;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
1 GGTTCGACGGCCCCCTGCTCAGCGGTGTATACATCCGCCCGGATACCAACCGGCCCC 59
|||
Db 106 GGCTCCACCGCCCCCGACGCCACGCGTGTACCTCGGCCCGGACACACGCGCGCCCC 164
|||

RESULT 3
BU542454

LOCUS BU542454 877 bp mRNA linear EST 13-SEP-2002
DEFINITION AGENCOURT_10322173 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574322
5', mRNA sequence.
BU542454
BU542454.1 GI:22852937

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1CM2769 row: h column: 02
High quality sequence stop: 760.
Location/Qualifiers

FEATURES
source
1.877
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6574322"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 40"

ORIGIN
Query Match 71.7%; Score 43; DB 5; Length 877;
Best Local Similarity 83.1%; Pred. No. 0.00059;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
1 GGTTCGACGGCCCCCTGCTCAGCGGTGTATACATCCGCCCGGATACCAACCGGCCCC 59
|||
Db 35 GGCTCCACCGCCCCCGACGCCACGCGTGTACCTCGGCCCGGACACACGCGCGCCCC 93
|||

RESULT 4 959 bp mRNA linear EST 14-NOV-2002
LOCUS CA489836
DEFINITION AGENCOURT_10810668 MAPcL Homo sapiens cDNA clone IMAGE:6722324 5',
mRNA sequence.
CA489836
CA489836.1 GI:24952627
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Kristi A. Egland, Ira Pastan
CDNA Library Preparation: Invitrogen Corp

[illegible]

```

COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: DCTD/DRP
              cDNA Library Preparation: Rubin Laboratory
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLCM2579 row: b column: 21
              High quality sequence stop: 177.
              Location/Qualifiers
                1. 1268
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:6384308"
                  /tissue_type="carcinoma, cell line"
                  /lab_host="DH10B (phage-resistant)"
                  /clone_lib="NIH MGC 40"
                  /note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
                  Site_2: EcoRI; cDNA made by oligo-dT priming.
                  Directionally cloned into EcoRI/XhoI sites using the
                  following 5' adaptor: GGCAcGAG(G). Library constructed by
                  Ling Hong in the laboratory of Gerald M. Rubin (University
                  of California, Berkeley) using ZAP-cDNA synthesis kit
                  (Stratagene) and Superscript II RT (Life Technologies).
                  Note: this is a NIH_MGC Library."

ORIGIN
Query Match      71.7%; Score 43; DB 5; Length 1268;
Best Local Similarity 83.1%; Pred. No. 0.0006;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY      1  GGTTGACGCGCCCCCTGCTCACGGGTAAATCCGCCCCGATACAGACCGGCCCC 59
        |||
        90  GGCTTCACCGCCCCCCAGCCCACGGTGTCACCTCGCCCCGAGACACAGCGCGCCCC 148

RESULT  9      BQ920055      1343 bp      mRNA      linear      EST 20-AUG-2002
BQ920055      AGENCOURT_10031674 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6481609
LOCUS      5', mRNA sequence.
ACCESSION      BQ920055
VERSION      BQ920055.1 GI:22334753
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
              1 (bases 1 to 1343)
              NIH-MGC http://mgs.nci.nih.gov/.
              National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: DCTD/DRP
              cDNA Library Preparation: Rubin Laboratory
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLCM2663 row: e column: 02
              High quality sequence start: 56
              High quality sequence stop: 237.
              Location/Qualifiers
                1. 1343
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:6481609"

FEATURES
source

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: January 15, 2005, 16:00:11 : Search time 170.5 Seconds
(without alignments)
1847.304 Million cell updates/sec

Title: US-10-057-136-7

Perfect score: 60
Sequence: 1 ggttcgacggccccccctgc.....cggataccagacggccct 60

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : N_Geneseq_23Sep04:*
- 1: _geneseqn1980s:*
 - 2: _geneseqn1990s:*
 - 3: _geneseqn2000s:*
 - 4: _geneseqn2001as:*
 - 5: _geneseqn2001bs:*
 - 6: _geneseqn2002as:*
 - 7: _geneseqn2002bs:*
 - 8: _geneseqn2003as:*
 - 9: _geneseqn2003bs:*
 - 10: _geneseqn2003cs:*
 - 11: _geneseqn2003ds:*
 - 12: _geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	60	100.0	60	2	AAV48320	AAV48320 Nucleotid
2	43	71.7	60	2	AAV48316	AAV48316 Nucleotid
C 3	43	71.7	309	1	AAN90579	Aan90579 pDF9.3 CD
4	43	71.7	1194	12	AD157712	Ad157712 Human bre
5	43	71.7	1378	12	AD157693	Ad157693 Human bre
C 6	43	71.7	1424	12	ADO23180	Ado23180 Antisense
7	43	71.7	1428	6	ABL60159	Ab160159 Human MUC
8	43	71.7	1428	12	ADO23125	Ado23125 Human MUC
9	43	71.7	1457	12	ADP32627	Adf32627 Plasmid J
10	43	71.7	1527	2	AAV48329	AAV48329 MiniMUC1
11	43	71.7	1614	12	ADK70370	Adk70370 Respirato
12	43	71.7	1630	12	AD157708	Ad157708 Human bre
13	43	71.7	1634	12	AD157689	Ad157689 Human bre
14	43	71.7	1712	12	AD157686	Ad157686 Human bre
15	43	71.7	1738	12	AD157669	Ad157669 Human bre
16	43	71.7	1755	12	AD157673	Ad157673 Human bre
17	43	71.7	1774	12	ADE43991	Ade43991 Plasmid J
18	43	71.7	1774	12	ADP32625	Adf32625 Plasmid J
19	43	71.7	1799	12	ADO23124	Ado23124 Human MUC
20	43	71.7	1803	12	AD157699	Ad157699 Human bre
21	43	71.7	1804	6	ABL67539	Ab167539 Thyroid c

22	43	71.7	1804	9	AAD56950	Ad56950 Human muc
23	43	71.7	1804	10	ADD14719	Add14719 Human src
24	43	71.7	1804	12	ADP13294	Adp13294 Renal cel
25	43	71.7	1805	12	ADO28642	Ado28642 Human MUC
26	43	71.7	1808	12	AD157706	Ad157706 Human bre
27	43	71.7	1818	12	ADP32633	Adf32633 Plasmid J
28	43	71.7	1823	6	ABZ35228	Abz35228 Human gen
29	43	71.7	1823	12	AD157707	Ad157707 Human bre
30	43	71.7	1835	12	ADP32631	Adf32631 Plasmid J
31	43	71.7	1874	12	AD157688	Ad157688 Human bre
32	43	71.7	1882	12	AD157677	Ad157677 Human bre
33	43	71.7	1918	12	AD157672	Ad157672 Human bre
34	43	71.7	1930	12	AD157678	Ad157678 Human bre
35	43	71.7	1945	12	AD157676	Ad157676 Human bre
36	43	71.7	1949	12	AD157698	Ad157698 Human bre
37	43	71.7	1953	12	AD157668	Ad157668 Human bre
38	43	71.7	2045	12	AD157701	Ad157701 Human bre
39	43	71.7	2049	12	AD157682	Ad157682 Human bre
40	43	71.7	2090	12	AD157705	Ad157705 Human bre
41	43	71.7	2094	12	AD157681	Ad157681 Human bre
42	43	71.7	2135	12	ADP32629	Adf32629 Plasmid J
43	43	71.7	2194	12	AD157683	Ad157683 Human bre
44	43	71.7	2255	12	AD157667	Ad157667 Human bre
45	43	71.7	2333	12	AD157685	Ad157685 Human bre

ALIGNMENTS

RESULT 1
ID AAV48320 standard; DNA; 60 BP.
XX AAV48320;
AC AAV48320;
XX 20-NOV-1998 (first entry)
DT
XX Nucleotide sequence encoding MUC1 tandem repeat unit R5.
DE
XX
XX
KW MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW tumour; tumour-associated antigen.
XX
OS Homo sapiens.
XX
XX
PN W09837095-A2.
XX
XX 27-AUG-1998.
PD
XX
XX 24-FEB-1998; 98WO-US003693.
PF
XX
XX 24-FEB-1997; 97US-0038253P.
PR
XX
XX (THER-) THERION BIOLOGICS CORP.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (DAND) DANA FARBER CANCER INST INC.
XX
PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX
XX WPI; 1998-467492/40.
PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
PS
XX Disclosure; Page 11; 42pp; English.
XX
XX The MUC1 tandem repeat units AAV48317-V48325 were used to create an
XX immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
XX (RPV). The RPV was used in a pharmaceutical composition also containing
XX an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
XX recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
XX does not undergo significant genetic deletion, thereby providing an
XX unexpected stable and immunogenic pox virus. They can be used to
XX prevent or treat tumours expressing MUC1 tumour-associated antigens

XX Sequence 60 BP; 9 A; 27 C; 15 G; 9 T; 0 U; 0 Other;
SQ

Query Match 100.0%; Score 60; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.9e-10;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGTTCGACGGCCCCCTGCTCAGGTGTAAATCCGCGGATACCAAGCCGCCCT 60
DB 1 GGTTCGACGGCCCCCTGCTCAGGTGTAAATCCGCGGATACCAAGCCGCCCT 60

RESULT 2

AAV48316
ID AAV48316 standard; cDNA; 60 BP.

AC AAV48316;

DT 20-NOV-1998 (first entry)

DE Nucleotide sequence encoding MUC1 tandem repeat unit.

KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW tumour; tumour-associated antigen.

OS Homo sapiens.

FH Key Location/Qualifiers
FT CDS 1..60
FT /*tag= a
FT /product= "MUC1 tandem repeat unit"

PN WO9837095-A2.

PD 27-AUG-1998.

PF 24-FEB-1998; 98WO-US003693.

PR 24-FEB-1997; 97US-0038253P.

PA (THER-) THERION BIOLOGICS CORP.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (DAND) DANA FARBER CANCER INST INC.

PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;

DR WPI; 1998-467492/40.

DR P-PSDB; AAW77229.

PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.

PS Example 1; Page 20; 42pp; English.

CC The MUC1 tandem repeat unit was used to create an immunogenic mini-MUC1
CC fragment for inclusion in a recombinant pox virus (RPV). The RPV was used
CC in a pharmaceutical composition also containing an immunomodulator to
CC generate MUC1 specific cytotoxic T-lymphocytes. The recombinant pox virus
CC therefore encodes an immunogenic MUC1 fragment that does not undergo
CC significant genetic deletion, thereby providing an unexpectedly stable
CC and immunogenic pox virus. They can be used to prevent or treat tumours
CC expressing MUC1 tumour-associated antigens

SQ Sequence 60 BP; 7 A; 33 C; 16 G; 4 T; 0 U; 0 Other;

Query Match 71.7%; Score 43; DB 2; Length 60;
Best Local Similarity 83.1%; Pred. No. 7.9e-05;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 GGTTCGACGGCCCCCTGCTCAGGTGTAAATCCGCGGATACCAAGCCGCCCT 59
DB 1 GGCTCCACCGCCCCCAGCCCAAGGTGTACCTCGGCCCGACACCAAGCCGCCCT 59

RESULT 3
AAN90579/c
ID AAN90579 standard; cDNA; 309 BP.

AC AAN90579;

DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 04-DEC-1989 (first entry)

DE PDF9.3 cDNA insert.

KW PDF9.3; human DF3 breast carcinoma-associated antigen epitope.

OS Human MCF-7 breast carcinoma cells; 'ATCC HTB22'.

PN WO8907107-A.

PD 10-AUG-1989.

PF 29-JAN-1988; 88US-00149831.

PR 29-JAN-1988; 88US-00149831.

PA (DANA-) DANA-FARBER CANCER.

PI Kufe DW;

DR WPI; 1989-248989/34.

DR P-PSDB; AAP91045, AAP91053, AAP91054, AAP91046.

PT Recombinant polypeptide(s) - contains DF 3 breast carcinoma antigen
PT epitope and useful as assay reagents, and encoding DNA sequences.

PS Claim 1; Fig 4; 31pp; English.

CC The sequence encodes a carbohydrate-free polypeptide contg. DF3 breast
CC carcinoma antigen epitope. Useful as a competitive binding assay reagent
CC and improves diagnosis. The cDNA consists of nearly identical 60 BP
CC tandem repeats which are 85% GC-rich. See also AAP91045, AAP91046,
CC AAP91053, AAP91054. (Updated on 25-MAR-2003 to correct PA field.)
CC (Updated on 27-AUG-2003 to correct OS field.)

SQ Sequence 309 BP; 21 A; 92 C; 160 G; 36 T; 0 U; 0 Other;

Query Match 71.7%; Score 43; DB 1; Length 309;
Best Local Similarity 83.1%; Pred. No. 8.9e-05;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 GGTTCGACGGCCCCCTGCTCAGGTGTAAATCCGCGGATACCAAGCCGCCCT 59
DB 266 GGCTCCACCGCCCCCAGCCCAAGGTGTACCTCGGCCCGACACCAAGCCGCCCT 208

RESULT 4

ADIS7712
ID ADIS7712 standard; cDNA; 1194 BP.

AC ADIS7712;

DT 22-APR-2004 (first entry)

DE Human breast specific nucleic acid (BSNA) #83.

KW Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
KW breast cancer; cytostatic.

OS Homo sapiens.

PN WO2003106648-A2.

PD 24-DEC-2003.

XX 16-JUN-2003; 2003WO-US018934.
XX
XX 14-JUN-2002; 2002US-0389327P.
XX
XX (DIAD-) DIADEXUS INC.
XX
PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
XX WPI; 2004-082185/08.
DR P-PSDB; ADI57782.
XX
PT Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.
XX
XX Claim 1; SEQ ID NO 83; 370pp; English.
XX
CC The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the
CC hybridisation. The nucleic acids are useful for determining the presence
CC of a BSP in a sample which involves contacting the sample with a BSP in
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein and
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.
XX
SQ Sequence 1194 BP; 243 A; 438 C; 275 G; 238 T; 0 U; 0 Other;

Query Match 71.7%; Score 43; DB 12; Length 1194;
Best Local Similarity 83.1%; Pred. No. 9.9e-05;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GGTTGACGGCCCCCTGCTCAGCGGTACATCCGCCCCGATACCAAGACCGGCCCC 59
Db 596 GGCTCCACCGCCCCCGCCAGCCACGGGTGTACCTCGGCCCCGACACAGCGCGGCC 654

RESULT 5
ADIS7693
ID ADIS7693 standard; cDNA; 1378 BP.
XX
AC ADIS7693;
XX
DT 22-APR-2004 (first entry)
XX
DE Human breast specific nucleic acid (BSNA) #64.
XX
KW Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
KM breast cancer; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2003106648-A2.
XX
PD 24-DEC-2003.
XX
PF 16-JUN-2003; 2003WO-US018934.
XX
PR 14-JUN-2002; 2002US-0389327P.

XX (DIAD-) DIADEXUS INC.
XX
XX Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
XX WPI; 2004-082185/08.
DR P-PSDB; ADI57765.
XX
PT Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.
XX
XX Claim 1; SEQ ID NO 64; 370pp; English.
XX
CC The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the
CC hybridisation. The nucleic acids are useful for determining the presence
CC of a BSP in a sample which involves contacting the sample with suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein and
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.
XX
SQ Sequence 1378 BP; 295 A; 489 C; 316 G; 278 T; 0 U; 0 Other;

Query Match 71.7%; Score 43; DB 12; Length 1378;
Best Local Similarity 83.1%; Pred. No. 0.0001;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GGTTGACGGCCCCCTGCTCAGCGGTACATCCGCCCCGATACCAAGACCGGCCCC 59
Db 596 GGCTCCACCGCCCCCGCCAGCCACGGGTGTACCTCGGCCCCGACACAGCGCGGCC 654

RESULT 6
ADO23180/c
ID ADO23180 standard; RNA; 1424 BP.
XX
AC ADO23180;
XX
DT 12-AUG-2004 (first entry)
XX
DE Antisense human MUC1 mucin glycoprotein RNA (coding sequence) SeqID 75.
XX
KW human; MUC1; mucin glycoprotein; cancer; chemotherapeutic; MUC1/ECF; ss;
KM antisense.
XX
OS Homo sapiens.
XX
PN WO2004044160-A2.
XX
PD 27-MAY-2004.
XX
PF 12-NOV-2003; 2003WO-US035848.
XX
PR 13-NOV-2002; 2002US-00293391.
XX
PR 29-MAY-2003; 2003US-00447839.
XX
PA (DAND) DANA FARBER CANCER INST INC.
PA (ILEX-) ILEX PROD INC.

```
XX PI Kufe DW, Kharbanda S, Weltman SD;
XX XX
DR WPI; 2004-420304/39.
XX XX
PT Double-stranded RNA complex useful for inhibiting proliferation of cancer
PT cell expressing MUC1 mucin glycoprotein, comprises first and second RNA
PT sequences.
XX XX
PS Disclosure; SEQ ID NO 75; 112pp; English.
XX XX
CC This invention relates to novel modulators of the human MUC1 mucin
CC glycoprotein for use in cancer therapeutics, where MUC1 is a protein that
CC acts to inhibit the apoptotic response to genotoxic stress caused by
CC chemotherapeutic agents. In particular, it refers to modulators of the
CC MUC1 extracellular domain (MUC1/ECD). The method refers to using double-
CC stranded RNA complexes as MUC1 interference RNA compositions such that
CC MUC1 expression is inhibited, which in turn inhibits cancer cell
CC proliferation. The present invention describes screening assays to
CC identify compounds that inhibit the binding of various MUC1 ligands such
CC as neueregulin 2, as well as agonists, antagonists and antibodies thereof.
CC Furthermore, it provides MUC1 antisense, small interfering RNA (siRNA)
CC and small molecules in combination with chemotherapeutic agents that are
CC useful in the field of cancer therapy. This polynucleotide sequence is
CC the antisense human MUC1 RNA of the invention.
XX XX
SQ Sequence 1424 BP; 309 A; 325 C; 495 G; 0 T; 295 U; 0 Other;

Query Match
Best Local Similarity 71.7%; Score 43; DB 12; Length 1424;
Matches 49; Conservative 83.1%; Pred. No. 0.0001;
Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTGACGCGCCCCCTGCTCAGCGGTGTACATCCGCCCGGATACCAAGCGGCCCC 59
Db 1041 GGCTCCACCGCCCCCGCCAGCCACGCGTGTCACTCGGCCCGGACACCAAGCGGCCCC 983

RESULT 7
ABL60159
ID ABL60159 standard; cDNA; 1428 BP.
XX XX
AC ABL60159;
XX XX
DT 22-JUL-2002 (first entry)
XX XX
DE Human MUC1 encoding cDNA SEQ ID NO 2.
XX XX
KW Human; mucin 1; MUC1; transmembrane protein; SNP; cancer; cytostatic;
KW single nucleotide polymorphism; haplotyping; genotyping; drug;
XX XX
KW antiinflammatory; gene; ss.
XX XX
OS Homo sapiens.
XX XX
FH Key
FH CDS Location/Qualifiers
FT 1..1428
FT /*tag= a
FT /product= "MUC1"
FT replace(1009,A)
FT /*tag= b
FT variation
FT /standard name= "Single nucleotide polymorphism"
FT /note= "SNP allelic variation results in Val substituted
FT by Met at position 337 of the MUC1 protein (ABB77476)"
XX XX
PN WO200226765-A2.
XX XX
PD 04-APR-2002.
XX XX
PF 25-SEP-2001; 2001WO-US030151.
XX XX
PR 28-SEP-2000; 2000US-0236113P.
XX XX
PA (GENA-) GENAISSANCE PHARM INC.
XX XX
```

```
PI Chew A, Koshy B;
XX XX
DR WPI; 2002-405042/43.
DR P-PSDB; ABB77476.
XX XX
PT New genetic variants of mucin 1, Transmembrane gene, useful in studying
PT expression and function of protein encoded by the gene and for screening
PT drugs to treat diseases e.g. cancer.
XX XX
PS Claim 23; Fig 2; 75pp; English.
XX XX
CC The invention relates to a polynucleotide (ABL60158, ABL60159) encoding
CC mucin 1/MUC1 (ABB77476), Transmembrane isogene. The invention describes
CC novel genetic variants of the MUC1 gene. The invention is useful for
CC haplotyping/genotyping the MUC1 gene in an individual and identifying an
CC association between a trait and at least one of the haplotypes or
CC haplotype pairs of MUC1 gene. MUC1 is useful for studying the expression
CC and function of MUC1 and expressing MUC1 protein for use in screening for
CC candidate drugs to treat diseases related to MUC1 activity and in
CC studying the effect of the variation on the biological activity of MUC1
CC as well as on the binding affinity of candidate drugs targeting MUC1 for
CC the treatment of e.g. cancer. MUC1 is further used by the pharmaceutical
CC research scientist to validate MUC1 as a candidate target for and in
CC design of clinical trials of candidate drugs for treating a specific
CC condition drugs or disease predicted to be associated with MUC1 activity.
CC MUC1 antibodies are useful in a variety of diagnostic and prognostic
CC formats and therapeutic methods
XX XX
SQ Sequence 1428 BP; 296 A; 496 C; 327 G; 309 T; 0 U; 0 Other;

Query Match
Best Local Similarity 71.7%; Score 43; DB 6; Length 1428;
Matches 49; Conservative 83.1%; Pred. No. 0.0001;
Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTGACGCGCCCCCTGCTCAGCGGTGTACATCCGCCCGGATACCAAGCGGCCCC 59
Db 385 GGCTCCACCGCCCCCGCCAGCCACGCGTGTCACTCGGCCCGGACACCAAGCGGCCCC 443

RESULT 8
ADO23125
ID ADO23125 standard; RNA; 1428 BP.
XX XX
AC ADO23125;
XX XX
DT 12-AUG-2004 (first entry)
XX XX
DE Human MUC1 mucin glycoprotein RNA (coding sequence) SegID 20.
XX XX
KW human; MUC1; mucin glycoprotein; cancer; chemotherapeutic; MUC1/ECD; ss.
XX XX
KW Homo sapiens.
XX XX
OS Homo sapiens.
XX XX
PN WO2004044160-A2.
XX XX
PD 27-MAY-2004.
XX XX
PF 12-NOV-2003; 2003WO-US035848.
XX XX
PR 13-NOV-2002; 2002US-00293391.
PR 29-MAY-2003; 2003US-00447839.
XX XX
PA (DAND ) DANA FARBER CANCER INST INC.
PA (ILEX-) ILEX PROD INC.
XX XX
PI Kufe DW, Kharbanda S, Weltman SD;
XX XX
DR WPI; 2004-420304/39.
XX XX
PT Double-stranded RNA complex useful for inhibiting proliferation of cancer
PT cell expressing MUC1 mucin glycoprotein, comprises first and second RNA
PT sequences.
XX XX
```

PS Claim 2; SEQ ID NO 20; 112pp; English.
XX
CC This invention relates to novel modulators of the human MUC1 mucin
CC glycoprotein for use in cancer therapeutics, where MUC1 is a protein that
CC acts to inhibit the apoptotic response to genotoxic stress caused by
CC chemotherapeutic agents. In particular, it refers to modulators of the
CC MUC1 extracellular domain (MUC1/EC). The method refers to using double-
CC stranded RNA complexes as MUC1 interference RNA compositions such that
CC MUC1 expression is inhibited, which in turn inhibits cancer cell
CC proliferation. The present invention describes screening assays to
CC identify compounds that inhibit the binding of various MUC1 ligands such
CC as neueregulin 2, as well as agonists, antagonists and antibodies thereof.
CC Furthermore, it provides MUC1 antisense, small interfering RNA (siRNA)
CC and small molecules in combination with chemotherapeutic agents that are
CC useful in the field of cancer therapy. This polynucleotide sequence is
CC the human MUC1 RNA of the invention.
XX
SQ Sequence 1428 BP; 296 A; 496 C; 327 G; 0 T; 309 U; 0 Other;

Query Match 71.7%; Score 43; DB 12; Length 1428;
Best Local Similarity 76.3%; Pred. No. 0.0001;
Matches 45; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 1 GGTTGACGGCCCCCTGCTCAGCGTGTACATCCGCCGATACCAAGCCGCCCC 59
DB 385 GGCTCCACCGCCCCCCAGCCCAAGCGUGUACACCCGCCGCCGAGACCAAGCCGCCCC 443

RESULT 9
ADF32627
ID ADF32627 standard; DNA; 1457 BP.
XX
AC ADF32627;
XX
DT 26-FEB-2004 (first entry)
XX
DE Plasmid JNW358 MUC-1 nucleotide sequence.
XX
KM MUC-1 antigen; immune response; MUC-1; variable number of tandem repeat;
KM VNTR; repeat unit; tumour; metastasis; cytostatic; vaccine; gene therapy;
KM gene; ds.
XX
OS Synthetic.
XX
PN WO2003100060-A2.
XX
PD 04-DEC-2003.
XX
PF 23-MAY-2003; 2003WO-EP005594.
XX
PR 24-MAY-2002; 2002GB-00012046.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Burden N, Ellis JH, Hamblin PA;
XX
DR WPI; 2004-042811/04.
XX
PT New nucleic acid molecule encoding a MUC-1 antigen, useful for preparing
PT a composition for treating or preventing tumors or metastases.
XX
PS Example; Fig 3; 66pp; English.
XX
CC The present invention describes a nucleic acid molecule which encodes a
CC MUC-1 antigen. The nucleic acid is capable of raising an immune response
CC in vivo, has reduced susceptibility to recombination than full-length MUC
CC -1 and comprises between 1 and 15 variable number of tandem repeats
CC (VNTR) perfect repeat units. Also described: (1) a plasmid comprising the
CC DNA molecule; (2) a protein encoded by the nucleic acid; (3) a
CC pharmaceutical composition comprising the nucleic acid, plasmid or
CC protein and an excipient, diluent or carrier; and (4) a method of
CC treating or preventing tumours or metastases. A MUC1 antigen has
CC cytostatic activity, and can be used in vaccines, and in gene therapy.

CC The nucleic acid is useful for preparing a composition for treating or
CC preventing tumours or metastases. The present sequence is used in the
CC exemplification of the present invention.
XX
SQ Sequence 1457 BP; 302 A; 506 C; 334 G; 315 T; 0 U; 0 Other;

Query Match 71.7%; Score 43; DB 12; Length 1457;
Best Local Similarity 83.1%; Pred. No. 0.0001;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GGTTGACGGCCCCCTGCTCAGCGTGTACATCCGCCGATACCAAGCCGCCCC 59
DB 403 GGCTCCACCGCCCCCCAGCCCAAGCGTGTACCTCGGCCCGGACACAGCGCGGCC 461

RESULT 10
AAV48329
ID AAV48329 standard; cDNA; 1527 BP.
XX
AC AAV48329;
XX
DT 20-NOV-1998 (first entry)
XX
DE MiniMUC1 gene.
XX
KM ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KM tumour; tumour-associated antigen.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1527
FT FT /*tag= a
FT FT /product= "MiniMUC1 protein"
XX
PN WO9837095-A2.
XX
PD 27-AUG-1998.
XX
PF 24-FEB-1998; 98WO-US003693.
XX
PR 24-FEB-1997; 97US-0038253P.
XX
PA (THER-) THERION BIOLOGICS CORP.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (DAND) DANA FARBER CANCER INST INC.
XX
PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX
DR WPI; 1998-467492/40.
DR P-PSDB; AAW77233.
XX
PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS Example 1; Page 21-22; 42pp; English.
XX
CC The immunogenic mini-MUC1 fragment was created from MUC1 tandem repeat
CC units for inclusion in a recombinant pox virus (RPV). The RPV was used in
CC a pharmaceutical composition also containing an immunomodulator to
CC generate MUC1 specific cytotoxic T-lymphocytes. The recombinant pox virus
CC therefore encodes an immunogenic MUC1 fragment that does not undergo
CC significant genetic deletion, thereby providing an unexpectedly stable
CC and immunogenic pox virus. They can be used to prevent or treat tumours
CC expressing MUC1 tumour-associated antigens
XX
SQ Sequence 1527 BP; 296 A; 573 C; 351 G; 307 T; 0 U; 0 Other;

Query Match 71.7%; Score 43; DB 2; Length 1527;
Best Local Similarity 83.1%; Pred. No. 0.0001;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GGTTGACGGCCCCCTGCTCAGCGTGTACATCCGCCCGGATACCAAGCCGCCCC 59

Db 226 GGCTCCACCGCCCCCGACCGGTGTCTACCTCGGCCCCGAGACACGAGCCGGCCCC 284

RESULT 11

ID ADK70370 standard; cDNA; 1614 BP.

AC ADK70370;

DT 06-MAY-2004 (first entry)

DE Respiratory disease differentially expressed cDNA #106.

XX ds; gene; cytostatic; respiratory; antiasthmatic; gene therapy;

KW differential gene expression; respiratory disorder; lung cancer;

KW chronic obstructive pulmonary disease; emphysema; asthma.

XX Homo sapiens.

XX WO2003101283-A2.

XX 11-DEC-2003.

PF 02-JUN-2003; 2003WO-US017409.

PR 04-JUN-2002; 2002US-0386005P.

PA (INCY-) INCYTE CORP.

PI Rickert PK, Krasnow R;

DR WPI; 2004-042945/04.

XX New combination comprising cDNAs and proteins that are differentially expressed in respiratory disorders, useful for diagnosing or treating respiratory diseases e.g. lung cancer, chronic obstructive pulmonary diseases or asthma.

PS Claim 1; SEQ ID NO 106; 343pp; English.

CC The invention relates to cDNA sequences that are differentially expressed in respiratory disorders or their complements or encoded proteins. The cDNAs and proteins are useful for diagnosing, treating or monitoring treatment of a subject with a respiratory disease including lung cancer, chronic obstructive pulmonary diseases, emphysema or asthma. The protein is also useful for screening molecules or compounds to identify at least one ligand which specifically binds the protein. It is also useful for preparing and purifying a polyclonal or monoclonal antibody. This sequence corresponds to a cDNA of the invention.

XX Sequence 1614 BP; 329 A; 556 C; 374 G; 353 T; 0 U; 2 Other;

Query Match 71.7%; Score 43; DB 12; Length 1614;

Best Local Similarity 83.1%; Pred. No. 0.0001;

Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GGTTGACGGCCCCCTGCTACCGGTGTACATCCGCCCGGATACGAGACCGGCCCC 59

Db 434 GGCTCCACCGCCCCCGACCGGTGTCTACCTCGGCCCCGAGACACGAGCCGGCCCC 492

RESULT 12

ID ADI57708 standard; cDNA; 1630 BP.

AC ADI57708;

DT 22-APR-2004 (first entry)

DE Human breast specific nucleic acid (BSNA) #79.

XX Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;

KW breast cancer; cytostatic.

XX Homo sapiens.

XX WO2003106648-A2.

XX 24-DEC-2003.

PF 16-JUN-2003; 2003WO-US018934.

PR 14-JUN-2002; 2002US-0389327P.

PA (DIAD-) DIADEXUS INC.

PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;

XX WPI; 2004-082185/08.

DR P-PSDB; ADI57779.

XX Novel isolated polypeptide comprising breast specific protein sequences, useful for diagnosing or monitoring presence and metastases of breast cancer in patient.

PS Claim 1; SEQ ID NO 79; 370pp; English.

CC The invention relates to human breast specific nucleic acids (BSNA) and the breast specific proteins (BSP) they encode. The nucleic acids are useful for determining the presence of a BSNA in a sample which involves contacting the sample with a BSNA under conditions in which the BSNA will selectively hybridise to a BSNA in the sample, and detecting the presence of a BSP in a sample which involves contacting the sample with suitable reagent under conditions in which the reagent will selectively interact with the BSP, and detecting the interaction of the reagent with a BSP in the sample. The nucleic acids and proteins are useful for diagnosing or monitoring the presence and metastases of breast cancer in a patient, which involves determining an amount of nucleic acid or protein and comparing the determined amount of nucleic acid or protein in the sample of the patient to the amount of a breast specific marker in a normal control, where a difference in the determined amount in the sample compared to the amount in the control is associated with the presence of breast cancer. The sequences are useful for treating a patient with breast cancer, involving administering a composition consisting of a BSNA or a BSP to a patient, where the administration induces an immune response against the breast cancer cell expressing the BSNA or BSP. This sequence represents a human BSNA of the invention.

XX Sequence 1630 BP; 322 A; 584 C; 405 G; 319 T; 0 U; 0 Other;

Query Match 71.7%; Score 43; DB 12; Length 1630;

Best Local Similarity 83.1%; Pred. No. 0.0001;

Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GGTTGACGGCCCCCTGCTACCGGTGTACATCCGCCCGGATACGAGACCGGCCCC 59

Db 596 GGCTCCACCGCCCCCGACCGGTGTCTACCTCGGCCCCGAGACACGAGCCGGCCCC 654

RESULT 13

ID ADI57689 standard; cDNA; 1634 BP.

AC ADI57689;

DT 22-APR-2004 (first entry)

DE Human breast specific nucleic acid (BSNA) #60.

XX Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;

KW breast cancer; cytostatic.

XX Homo sapiens.


```
PN WO2003106648-A2.
XX
XX 24-DEC-2003.
PD
XX
XX 16-JUN-2003; 2003WO-US018934.
PF
XX
XX 14-JUN-2002; 2002US-0389327P.
PR
XX
XX (DIAD-) DIADEXUS INC.
PA
XX
XX Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
PI
XX WPI; 2004-082185/08.
XX
XX P-PSDB; ADI57761.
DR
XX
XX Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.
XX
XX Claim 1; SEQ ID NO 60; 370pp; English.
XX
XX The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the
CC hybridisation. The nucleic acids are useful for determining the presence
CC of a BSP in a sample which involves contacting the sample with suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein and
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.
XX
XX Sequence 1634 BP; 349 A; 541 C; 405 G; 337 T; 0 U; 2 Other;
SQ
Query Match 71.7%; Score 43; DB 12; Length 1634;
Best Local Similarity 83.1%; Pred. No. 0.0001;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 GGTTCGACGGCCCCCTGCTCAGCGTGTACATCCGCCCGGATACCAAGCCGGCCCC 59
Db 596 GGCTTCACCGCCCCCCAGCCACCGGTGTCACTCGGCCCGGACACAGCGCCGCCCC 654
RESULT 14
ADIS7686
ID ADIS7686 standard; cDNA; 1712 BP.
XX
XX ADIS7686;
AC
XX 22-APR-2004 (first entry)
DT
XX Human breast specific nucleic acid (BSNA) #57.
DE
XX Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
KM breast cancer; cytostatic.
XX
XX Homo sapiens.
OS
XX WO2003106648-A2.
PN
XX 24-DEC-2003.
PD
XX
```

```
PF 16-JUN-2003; 2003WO-US018934.
XX
XX 14-JUN-2002; 2002US-0389327P.
PR
XX
XX (DIAD-) DIADEXUS INC.
PA
XX
XX Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
PI
XX WPI; 2004-082185/08.
XX
XX P-PSDB; ADI57758.
DR
XX
XX Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.
XX
XX Claim 1; SEQ ID NO 57; 370pp; English.
XX
XX The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the
CC hybridisation. The nucleic acids are useful for determining the presence
CC of a BSP in a sample which involves contacting the sample with suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein and
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.
XX
XX Sequence 1712 BP; 327 A; 604 C; 435 G; 346 T; 0 U; 0 Other;
SQ
Query Match 71.7%; Score 43; DB 12; Length 1712;
Best Local Similarity 83.1%; Pred. No. 0.0001;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 GGTTCGACGGCCCCCTGCTCAGCGTGTACATCCGCCCGGATACCAAGCCGGCCCC 59
Db 596 GGCTTCACCGCCCCCCAGCCACCGGTGTCACTCGGCCCGGACACCAAGCCGCCCC 654
RESULT 15
ADIS7669
ID ADIS7669 standard; cDNA; 1738 BP.
XX
XX ADIS7669;
AC
XX 22-APR-2004 (first entry)
DT
XX Human breast specific nucleic acid (BSNA) #40.
DE
XX Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
KM breast cancer; cytostatic.
XX
XX Homo sapiens.
OS
XX WO2003106648-A2.
PN
XX 24-DEC-2003.
PD
XX 16-JUN-2003; 2003WO-US018934.
PF
XX 14-JUN-2002; 2002US-0389327P.
PR
XX
```

PA (DIAD-) DIADEXUS INC.
XX
PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
XX
DR MPI; 2004-082185/08.
DR P-PSDB; ADI57743.
XX
PT Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.
XX
PS Claim 1; SEQ ID NO 40; 370pp; English.
XX
CC The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the
CC hybridisation. The nucleic acids are useful for determining the presence
CC of a BSP in a sample which involves contacting the sample with suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein and
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.
XX
SQ Sequence 1738 BP; 342 A; 642 C; 399 G; 355 T; 0 U; 0 Other;

Query Match 71.7%; Score 43; DB 12; Length 1738;
Best Local Similarity 83.1%; Pred. No. 0.0001;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GGTCGACGGCCCCCTGCTCAGCGGTGTACATCCGCGCGGATACAGACCGGCCCC 59
|||
Db 458 GGCTCCACCGCCCCCGCCAGCCACCGGTGTACCTCGGCCCCGGACACAGCGCGCCCC 516
|||

Search completed: January 15, 2005, 20:36:09
Job time : 170.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 17:40:52 ; Search time 722.8 Seconds
(without alignments)
3925.542 Million cell updates/sec

Title: US-10-057-136-7

Perfect score: 60
Sequence: 1 ggttcgacgagccccctgc.....cggataccagaccggccct 60

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	2297	6	BD272907 A recombi
2	43	71.7	120	6	AX192396 Sequence
3	43	71.7	120	6	BD000571 Human pol
4	43	71.7	1414	12	AF423031 Synthetic
5	43	71.7	1455	6	CQ715242 Sequence
6	43	71.7	1457	6	AX959914 Sequence
7	43	71.7	1774	6	AX959684 Sequence
8	43	71.7	1774	6	AX959912 Sequence
9	43	71.7	1804	6	AR492318 Sequence
10	43	71.7	1804	6	AX335367 Sequence
11	43	71.7	1804	9	HUMMUCAB
12	43	71.7	1834	12	AF423030 Synthetic
13	43	71.7	1835	6	AX959918 Sequence
14	43	71.7	2135	6	AX959916 Sequence
15	43	71.7	4139	6	CQ834017 Sequence
16	43	71.7	4139	6	AX334899 Sequence
17	43	71.7	4139	6	AX335372 Sequence
18	43	71.7	4139	6	AX336712 Sequence
19	43	71.7	4139	6	AX409474 Sequence

20	43	71.7	4139	6	AX440481 Sequence
21	43	71.7	4139	9	HUMPANMU
22	43	71.7	7188	9	AY463543 Homo sapi
23	43	71.7	8181	6	AX406624 Sequence
24	43	71.7	8181	9	HUMPEM
25	43	71.7	8186	6	AR492320 Sequence
26	43	71.7	133525	9	AL713999 Human DNA
27	41.4	69.0	120	6	AX192397 Sequence
28	41.4	69.0	120	6	BD000572 Human pol
29	41.4	69.0	180	6	E08763 CDNA encodi
30	41.4	69.0	572	6	AR492319 Sequence
31	41.4	69.0	572	9	HUMDF3A
32	41.4	69.0	1572	6	AX093798 Sequence
33	41.4	69.0	1721	6	CQ771290 Sequence
34	41.4	69.0	1721	6	AR492306 Sequence
35	41.4	69.0	1721	6	AX335860 Sequence
36	41.4	69.0	1721	6	AX440427 Sequence
37	41.4	69.0	1721	6	AX587588 Sequence
38	41.4	69.0	1721	9	HSTEYMA
39	41.4	69.0	1971	6	AX963157 Sequence
40	41.4	69.0	2037	6	AX963159 Sequence
41	41.4	69.0	2238	9	HSSETA
42	41	68.3	491	9	HUMEPISB1
43	41	68.3	518	6	AR492402 Sequence
44	41	68.3	518	9	HUMEPIS1A1
45	39.8	66.3	3343	6	AR492403 Sequence

ALIGNMENTS

RESULT 1
BD272907 2297 bp DNA linear PAT 17-JUL-2003
LOCUS A recombinant vector expressing multiple constitutimulatory molecules
DEFINITION and uses thereof.
ACCESSION BD272907 GI:33082675
VERSION JP 2002531133-A/1.
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 2297)
AUTHORS Schlom,J., Hodge,J. and Panicali,D.
TITLE A recombinant vector expressing multiple constitutimulatory molecules and uses thereof
JOURNAL Patent: JP 2002531133-A 1 24-SEP-2002;
COMMENT THE UNITED STATES OF AMERICA, THERION BIOLOGICS CORP
OS Artificial Sequence
PN JP 2002531133-A/1
PD 24-SEP-2002
PF 12-NOV-1999 JP 2000586927
PR 09-DEC-1998 US 60/111582
PI JEFFREY SCHLOM, JAMES HODGE, DENNIS PANICALI
PC C12N15/02, A61K35/12, A61K35/74, A61K35/76, A61K38/00, A61K39/00, A61K39/12, A61K39/21, A61K39/235, A61K39/245, A61K39/275, A61K39/29, A61K48/00, A61P1/04, A61P31/04, A61P31/10, A61P31/12, A61P35/00, A61P37/02, PC A61P37/06, A61P37/15, C12N1/19, C12N1/21, C12N5/10, C12N7/00, C12Q1/02, G01N33/53, C12N15/00, C12N5/00, A61K37/02
PC Description of Artificial Sequence: VECTOR SEQUENCE FH Key
CC Location/Qualifiers
FT source 1. .2297
FT location/Qualifiers
FEATURES
source 1. .2297
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
ORIGIN

Query Match 100.0%; Score 60; DB 6; Length 2297;
Best Local Similarity 100.0%; Pred. No. 5.8e-08;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCGACGGCCCCCTGCTCAGCGGTGAACATCCGCGGATACGACCGGCCCT 60
|||||
586 GGCTCCACGGCCCCCTGCTCAGCGGTGAACATCCGCGGATACGACCGGCCCT 645

RESULT 2
AXI92396/c AXI92396 120 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 2 from Patent EP1103623.
ACCESSION AXI92396
VERSION AXI92396.1 GI:15210363
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Taylor-Papadimitriou,J., Burchell,J. and Gendler,S.
TITLE Human mucin core protein: nucleic acid probes, peptide fragments
and antibodies thereto, and uses thereof in diagnostic and
therapeutic methods
Patent: EP 1103623-A 2 30-MAY-2001;
JOURNAL IMPERIAL CANCER RESEARCH TECHNOLOGY LIMITED (GB)
FEATURES location/Qualifiers
source 1. .120
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 71.7%; Score 43; DB 6; Length 120;
Best Local Similarity 83.1%; Pred. No. 0.014;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GGTTCGACGGCCCCCTGCTCAGCGGTGAACATCCGCGGATACGACCGGCC 59
|||||
87 GGCTCCACGGCCCCCTGCTCAGCGGTGAACATCCGCGGATACGACCGGCC 29

Db

RESULT 3
BD000571/c BD000571 120 bp DNA linear PAT 31-JAN-2002
DEFINITION Human polymorphic epithelial mucin core protein and nucleic acid
encoding the protein.
ACCESSION BD000571
VERSION BD000571.1 GI:18623684
KEYWORDS JP 2000333675-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 120)
Papadimitrov,J.T., Jendora,S. and Bachieru,J.
Human polymorphic epithelial mucin core protein and nucleic acid
encoding the protein
Patent: JP 2000333675-A 2 05-DEC-2000;
JOURNAL IMPERIAL CANCER RESEARCH TECHNOLOGY LTD
OS Homo sapiens (human)
PN JP 2000333675-A/2
PD 05-DEC-2000
PF 26-APR-2000 JP 2000125724
PR 07-JAN-1987 GB 8700269,07-JAN-1987 GB 8700279 PR
22-APR-1987 US 041306,09-NOV-1987 GB 8726172 PI JOYCE
TAYLOR PAPADIMITROV,SANDRA JENDORA,JOY BACHIERU PC
C12N15/02,A61K38/00,A61K39/395,A61K39/395,A61K49/00,A61P35/00, PC
C07K14/47,
PC C07K16/44,C12N5/10,C12P21/08// (C12N15/02,C12R1:91), (C12N5/10,
PC C12R1:91),

PC C12N15/00,A61K37/02,C12N5/00,(C12N15/00,C12R1:91), (C12N5/00,
PC C12R1:91)
CC
FH Key location/Qualifiers
FT source 1. .120
/organism="Homo sapiens (human)".
FEATURES source 1. .120
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 71.7%; Score 43; DB 6; Length 120;
Best Local Similarity 83.1%; Pred. No. 0.014;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GGTTCGACGGCCCCCTGCTCAGCGGTGAACATCCGCGGATACGACCGGCC 59
|||||
87 GGCTCCACGGCCCCCTGCTCAGCGGTGAACATCCGCGGATACGACCGGCC 29

Db

RESULT 4
AF423031 1414 bp mRNA linear SYN 10-JUL-2003
LOCUS AF423031
DEFINITION Synthetic construct Homo sapiens mucin variant MUC1-CT58 (MUC1)
mRNA, complete cds; alternatively spliced.
ACCESSION AF423031
VERSION AF423031.1 GI:19338621
KEYWORDS
SOURCE synthetic construct
synthetic construct
artificial sequences.
1 (bases 1 to 1414)
REFERENCE 1
AUTHORS Hinojosa-Kurtzberg,A.M., Johansson,M.E., Madsen,C.S., Hansson,G.C.
and Gendler,S.J.
TITLE Novel MUC1 splice variants contribute to mucin overexpression in
CFTR-deficient mice
JOURNAL Am. J. Physiol. Gastrointest. Liver Physiol. 284 (5), G853-G862
(2003)
MEDLINE 22570517
PUBMED 12529261
REFERENCE 2 (bases 1 to 1414)
AUTHORS Hinojosa-Kurtzberg,A.M. and Gendler,S.J.
TITLE Direct Submission
JOURNAL Submitted (21-SEP-2001) Biochemistry, Mayo Clinic, 13400 E. Shea
Boulevard, Scottsdale, AZ 85259, USA
FEATURES location/Qualifiers
source 1. .1414
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/note="Homo sapiens gene in transgenic Mus musculus
C57BL/6; isolated from intestinal mucosa"
1. .1414
/gene="MUC1"
1. .1386
/gene="MUC1"
/note="alternatively spliced; contains exon 6b resulting
in variant carboxy-terminal domain; lacks sites for
beta-catenin and Grb2 interactions; derived from Homo
sapiens"
/codon_start=1
/transl_table=1
/product="mucin variant MUC1-CT58"
/protein_id="AAU86735.1"
/db_xref="GI:19338622"
/translation="MTPGTQSPFFLLLLTLVLTVTGSGHASTPGGKETSATORSS
VPSSTEKNAVSMSTSVLSHSGSGSSTTGODVTIAPAREPASGAATWGQDVTSV
VTRPALGSTTPPAHDVTSAPDNKPAPGSTAPPAHGVTSADPTRPAGSTAPPAHGVTS
APDNRPALGSTAPVHNVTASAGSASTLVNNGTSARATTPASKSTPESIPSH
SDPTTLASHSTKTDASSTHSTVPPPLTSGSNHSTSPQLSTGVSFFLSFHISNLQFNS
SLSDPSTDYQELQRDISEMFLQIYKQGFLLGINSIKRPPSGSVVVLTLAFREGTIV

RESULT 9
AR492318

LOCUS	AR492318	1804 bp	DNA	linear	PAT 15-MAY-2004
DEFINITION	Sequence 17 from patent US 6716627.				
ACCESSION	AR492318				
VERSION	AR492318.1	GI:47260892			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 1804)				
TITLE	Dobie,K.W.				
JOURNAL	Antisense modulation of mucin 1, transmembrane expression				
FEATURES	Patent: US 6716627-A 17 06-APR-2004;				
source	Location/Qualifiers				
	1..1804				
	/organism="unknown"				
	/mol_type="genomic DNA"				
ORIGIN					
Query Match	71.7%;	Score 43;	DB 6;	Length 1804;	
Best Local Similarity	83.1%;	Pred. No. 0.0078;			
Matches	49;	Conservative	0;	Mismatches 10;	Indels 0;
Gaps					0;
QY	1	GGTTGACGCGCCCCCTGCTCAGCGGTGTAACATCCGCCCCCGGATACCAAGACGGCCCC	59		
Db	457	GGCTCCACCGCCCCCCCGACCGCCAGCGGTGTACCTTCGCCCCCGGACACCAAGCGCGCCCC	515		
RESULT 10					
AX335367					
LOCUS	AX335367	1804 bp	DNA	linear	PAT 09-JAN-2002
DEFINITION	Sequence 5876 from Patent WO0194629.				
ACCESSION	AX335367				
VERSION	AX335367.1	GI:18126086			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	1 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,				
JOURNAL	Hortigan,S., Soppet,D.R. and Weaver,Z.				
FEATURES	Cancer gene determination and therapeutic screening using signature				
source	gene sets				
	Patent: WO 0194629-A 5876 13-DEC-2001;				
	Avalon Pharmaceuticals (US)				
	Location/Qualifiers				
	1..1804				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
ORIGIN					
Query Match	71.7%;	Score 43;	DB 6;	Length 1804;	
Best Local Similarity	83.1%;	Pred. No. 0.0078;			
Matches	49;	Conservative	0;	Mismatches 10;	Indels 0;
Gaps					0;
QY	1	GCTTGACGCGCCCCCTGCTCAGCGGTGTAACATCCGCCCCCGGATACCAAGACGGCCCC	59		
Db	457	GGCTCCACCGCCCCCCCGACCGCCAGCGGTGTACCTTCGCCCCCGGACACCAAGCGCGCCCC	515		
RESULT 11					
HUMMUCAB					
LOCUS	HUMMUCAB	1804 bp	mRNA	linear	PRI 07-JAN-1995
DEFINITION	Human polymorphic epithelial mucin (PEM)				
ACCESSION	J05581				
VERSION	J05581.1	GI:188869			
KEYWORDS	polymorphic epithelial mucin.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
ORIGIN	1 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
	(bases 1 to 1804)				

AUTHORS	Gendler, S.J., Lancaster, C.A., Taylor-Papadimitriou, J., Duhig, T., Peat, N., Burchell, J., Pemberton, L., Lalani, E.N. and Wilson, D.			
TITLE	Molecular cloning and expression of human tumor-associated polymorphic epithelial mucin			
JOURNAL	J. Biol. Chem. 265 (25), 15286-15293 (1990)			
MEDLINE	90368715			
PUBMED	1697589			
COMMENT	Original source text: Homo sapiens adult adenocarcinoma cDNA to mRNA. Draft entry and computer-readable sequence for [J. Biol. Chem. (1990) In press] kindly submitted by S.J.Gendler, 26-JUN-1990.			
FEATURES	Location/Qualifiers			
source	1. 1804			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/map="Xq26.3-q27.1"			
	/cell_line="BT20"			
	/tissue_type="adenocarcinoma"			
	/dev_stage="adult"			
	1. 1804			
gene	/gene="MCNAA"			
	<1. 1804			
mRNA	/gene="MCNAA"			
	/product="PEM mRNA"			
CDS	73. 1500			
	/gene="MCNAA"			
	/codon_start=1			
	/product="polymorphic epithelial mucin"			
	/protein_id="AA59876.1"			
	/db_xref="GI:188870"			
	/translation="MTPGTOSPFLILLTLVTVVSGHASTPGEKETSATQRSSVPSSTEKNVAVMTSSVLSSHPGSGSTTGGQDVTLPATEPAGSAATWGQVTSVPVTRPALGSTTPPAHDVTSAPDNKPAGSTAPPAHGVTSAPDTRAPAGVTSAPDNRPALGSTAPVPAHNVTSAGSAGSASTLVANGTSARATTPASKSTPFSIPISSHSDPTTILASHSTKIDASTHSTVPLTSSNHSTSPOLSTGVSEFFLSFHSINQENS LLEDPTDYVOELORDISEMELQIKGGFLGLSNIKFRPGSVVQTLAFREGTINVHDVETQENQYKTEAASRYNLTISDVSVDPFPESAQSGAGVGMGIALVLVCLVA LAIVYLIALAVCCGCRKNYGOLDIFPARDTYHPMSEYPTYHTHGRYVPPSSITDRSPYE KVSAGNGGSSLSYTNPAVAATSANL"			
	73. 135			
sig_peptide	/gene="MCNAA"			
	136. 1497			
mat_peptide	/gene="MCNAA"			
	/product="polymorphic epithelial mucin"			
	1783. 1788			
polyA_signal	/gene="MCNAA"			
ORIGIN				
Query Match	71.7%; Score 43; DB 9; Length 1804;			
Best Local Similarity	83.1%; Pred. No. 0.0078;			
Matches 49; Conservative	0; Mismatches 10; Indels 0; Gaps 0;			
QY	1 GGTTCGACGGCCCCCTGCTCAGGTTGAACATCCGCCCGGATACGACGGCCCC 59			
Db	457 GGCCTCACCGCCCCCCAGCCGCGTGTCACTCGCCCCGAGACACGAGCGGCCCC 515			
RESULT 12				
AF423030				
LOCUS	1834 bp mRNA linear SYN 10-JUL-2003			
DEFINITION	Synthetic construct Homo sapiens mucin variant MUC-1CT80 (MUC1)			
ACCESSION	AF423030			
VERSION	AF423030.1 GI:19338619			
KEYWORDS	-			
SOURCE	synthetic construct			
ORGANISM	synthetic construct			
REFERENCE	artificial sequences.			
AUTHORS	1 (bases 1 to 1834)			
	Hinojosa-Kurtzberg, A.M., Johansson, M.E., Madgen, C.S., Hansson, G.C. and Gendler, S.J.			

FEATURES	source
TITLE	Novel MUC1 splice variants contribute to mucin overexpression in CFTR-deficient mice
JOURNAL	Am. J. Physiol. Gastrointest. Liver Physiol. 284 (5), G853-G862 (2003)
MEDLINE	22570517
PUBMED	12529261
REFERENCE	2 (bases 1 to 1834)
AUTHORS	Hinojosa-Kurtzberg, A.M. and Gendler, S.J.
TITLE	Direct Submission
JOURNAL	Submitted (21-SEP-2001) Biochemistry, Mayo Clinic, 13400 E. Shea Boulevard, Scottsdale, AZ 85259, USA
FEATURES	Location/Qualifiers
source	1. 1834
	/organism="synthetic construct"
	/mol_type="mRNA"
	/db_xref="taxon:32630"
gene	/note="Homo sapiens gene in transgenic Mus musculus C57BL/6; isolated from intestinal mucosa"
	1. 1834
	/gene="MUC1"
CDS	1. 1452
	/gene="MUC1"
	/note="alternatively spliced; contains exon 6a resulting in variant hydrophilic tail of carboxy-terminal domain; lacks sites for beta-catenin and Gtb2 interactions; derived from Homo sapiens"
	/codon_start=1
	/transl_table=11
	/product="mucin variant MUC-1CT80"
	/protein_id="AAU86734.1"
	/db_xref="GI:19338620"
	/translation="MTPGTOSPFLLLLTVLTVTGSGHASTPGGEKETSA TORSS VPSSTEKNVAVSMSTSVLSHSPGSGSSTGGQDVTLAPATEPASGSAATWGQVTSVP VTRPALGSTTPPAHDVTSAPDNKPAAGSTAPAHGVTAPDTRPAPGSTAPAHGVTSS APDNRPALGSTAPPAHNVTSASGSASGASTLVHNGTSARATTPASKSTPFSIPSHH SDPTTLASHSTKTSDASTHSTVPLTSSNHSTSPQLSTGVSFFLSFHISNLQFNS SLSDPSTDYQELQRDISEMFLQIYKGGFLGLSNKFRPGSVVQTLAFREGTINW HDVETQFNQYKTEAASRYNLTISDVSVDVPPFPFASQSGAVPGMGIALLVLCVLANV LAIVYLIALLAVCCQRRKNYGQLDIFPARDTYHPMSERYPTYHTGRVYPPSSIDRSPYE KVRLEPTGGKQKRVMLGKDESGGTWKTQRAWKR"
ORIGIN	
Query Match	71.7%; Score 43; DB 12; Length 1834;
Best Local Similarity	83.1%; Pred. No. 0.0078;
Matches	49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Db	385 GGCTCCACCGCCCCCGACCGGTGTACCTCGGCCCGGACACCGCGGGCCCC 443
RESULT 13	
LOCUS	AX959918
DEFINITION	AX959918 1835 bp DNA linear PAT 14-JAN-2004
ACCESSION	AX959918
VERSION	AX959918.1 GI:40880145
KEYWORDS	
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE	1
AUTHORS	Burden, N.G., Ellis, J.H. and Hamblin, P.A.
TITLE	Muc-1 antigen with reduced number of vntx repeat units
JOURNAL	Patent: WO 03100060-A 23 04-DEC-2003;
FEATURES	GLAXO GROUP LIMITED (GB)
source	Location/Qualifiers
	1. 1835
	/organism="Homo sapiens"
	/mol_type="unassigned DNA"
	/db_xref="taxon:9606"

Query Match	71.7%;	Score 43;	DB 6;	Length 1835;
Best Local Similarity	83.1%;	Pred. No. 0.0078;		
Matches 49;	Conservative 0;	Mismatches 10;	Indels 0;	Gaps 0;
Db	462	GGCTCCACCGCCCCCCCCCAGCCACCGGTGTCTCCTCGCCCCCGACACACGCGGCC	520	
RESULT 14				
LOCUS	AX959916	2135 bp	DNA	PAT 14-JAN-2004
DEFINITION	Sequence 21 from Patent WO03100060.			
ACCESSION	AX959916			
VERSION	AX959916.1	GI:40880144		
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	1 Burden,N.G., Ellis,J.H. and Hamblin,P.A.			
AUTHORS	Muc-1 antigen with reduced number of vnter repeat units			
TITLE	Patent: WO 03100060-A 21 04-DEC-2003;			
JOURNAL	GLAXO GROUP LIMITED (GB)			
FEATURES	Location/Qualifiers			
SOURCE	1..2135			
	/organism="Homo sapiens"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:9606"			
ORIGIN				
Query Match	71.7%;	Score 43;	DB 6;	Length 2135;
Best Local Similarity	83.1%;	Pred. No. 0.0075;		
Matches 49;	Conservative 0;	Mismatches 10;	Indels 0;	Gaps 0;
Db	462	GGCTCCACCGCCCCCCCCCAGCCACCGGTGTCTCCTCGCCCCCGACACACGCGGCC	520	
RESULT 15				
LOCUS	CQ834017	4139 bp	DNA	linear PAT 29-JUL-2004
DEFINITION	Sequence 53 from Patent EP1439393.			
ACCESSION	CQ834017			
VERSION	CQ834017.1	GI:50833622		
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	1 Astle,J.H., Boardman,L.A., Bugart,L.J., Burgess,C.C., Catino,T.J.,			
AUTHORS	Dwivedi,P., Huntress,M., Johnson,K.A., Lewis,M.E., Maimonis,P.J.,			
	Myerow,S.H., Brown-Shimer,S.L., Thiagalingam,A., Thibodeau,S.N. and			
	Molino,G.A.			
	Detection methods using TIMP 1 for colon cancer diagnosis			
	Patent: EP 1439393-A 53 21-JUL-2004;			
	Bayer Healthcare LLC (US); MAYO FOUNDATION FOR MEDICAL EDUCATION			
	AND RESEARCH (US)			
FEATURES	Location/Qualifiers			
SOURCE	1..4139			
	/organism="Homo sapiens"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:9606"			
ORIGIN				
Query Match	71.7%;	Score 43;	DB 6;	Length 4139;
Best Local Similarity	83.1%;	Pred. No. 0.0065;		
Matches 49;	Conservative 0;	Mismatches 10;	Indels 0;	Gaps 0;

QY 1 GGTTGACGGCCCCCCTGCTCAGCGTGTACATCCGCCCCGGATACAGACCGGCCCC 59
Db 458 GGCTCCACCGCCCCCAGCCACGCGTGTCACTCGGCCCCGGACACCAAGCCGGCCCC 516

Search completed: January 15, 2005, 22:36:52
Job time : 722.8 secs


```
Db      1 GGGTCAACAGCTCCTCCCGCTCATGGGGTTACTTCTGCTCCAGATACTCGCCAGCTCCA 60
RESULT 2
US-10-406-317-41
; Sequence 41, Application US/10406317
; Publication No. US20040019195A1
; GENERAL INFORMATION:
; APPLICANT: Schlom, Jeffrey;
; APPLICANT: Hodge, James;
; APPLICANT: Panicali, Dennis
; TITLE OF INVENTION: A recombinant vector expressing multiple constimulatory
; FILE REFERENCE: 38163-0189
; CURRENT APPLICATION NUMBER: US/10/406,317
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US/09/856,988
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: PCT/US99/26866
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/111,582
; PRIOR FILING DATE: 1998-12-09
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 41
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: VECTOR
US-10-406-317-41

Query Match      100.0%; Score 60; DB 16; Length 2297;
Best Local Similarity 100.0%; Pred. No. 7.7e-12;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGGTCAACAGCTCCTCCCGCTCATGGGGTTACTTCTGCTCCAGATACTCGCCAGCTCCA 60
Db      526 GGGTCAACAGCTCCTCCCGCTCATGGGGTTACTTCTGCTCCAGATACTCGCCAGCTCCA 585

RESULT 3
US-10-057-136-11
; Sequence 11, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 11
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-11

Query Match      61.0%; Score 36.6; DB 14; Length 60;
Best Local Similarity 76.3%; Pred. No. 0.0014;
```

```
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY      1 GGGTCAACAGCTCCTCCCGCTCATGGGGTTACTTCTGCTCCAGATACTCGCCAGCTCC 59
Db      1 GGTTCACGCACTCTCCAGACACGAGTCACTGTGACCCGACACCCGCTCAGCTCC 59

RESULT 4
US-10-057-136-10
; Sequence 10, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 10
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-10

Query Match      60.0%; Score 36; DB 14; Length 60;
Best Local Similarity 75.0%; Pred. No. 0.0023;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY      1 GGGTCAACAGCTCCTCCCGCTCATGGGGTTACTTCTGCTCCAGATACTCGCCAGCTCCA 60
Db      1 GGGTCACTGCCCCCTCCGCGCATGTGTGACCTCAGCTCCTGACACAGGCCAGCCCA 60

RESULT 5
US-10-057-136-5
; Sequence 5, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 5
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-5
```


Query Match 49.3%; Score 29.6; DB 17; Length 93;
Best Local Similarity 79.5%; Pred. No. 0.5;
Matches 35; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 1 GGGTCAACAGCTCTCCCGCTCATGGGTTACTTCTGCTCCAGA 44
|||
Db 57 GGCTCAACAGCCCCCGCAGCTCATGTGTCACTCAGCTCCCGA 14

RESULT 10

US-10-471-607-6
; Sequence 6, Application US/10471607
; Publication No. US20040115740A1
; GENERAL INFORMATION:
; APPLICANT: The Victoria University of Manchester
; APPLICANT: Benson, Roderick
; TITLE OF INVENTION: Intracellular analysis.
; FILE REFERENCE: P088857PMO
; CURRENT APPLICATION NUMBER: US/10/471,607
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: GB 0108165.2
; PRIOR FILING DATE: 3001-03-21
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 157
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Artificial epitope construct
US-10-471-607-6

Query Match 49.3%; Score 29.6; DB 17; Length 157;
Best Local Similarity 79.5%; Pred. No. 0.54;
Matches 35; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 1 GGGTCAACAGCTCTCCCGCTCATGGGTTACTTCTGCTCCAGA 44
|||
Db 101 GGCTCAACAGCCCCCGCAGCTCATGTGTCACTCAGCTCCCGA 144

RESULT 11

US-10-057-136-12
; Sequence 12, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALLI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; TITLE OF INVENTION: TUMOR-ASSOCIATED ANTIGEN
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-12

Query Match 48.3%; Score 29; DB 14; Length 60;
Best Local Similarity 71.7%; Pred. No. 0.77;

Matches 38; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY 7 ACAGCTCTCCCGCTCATGGGTTACTTCTGCTCCAGATACTGCCAGCTCC 59
|||
Db 7 ACAGCGCCACCCGACATAGGCTCACGAGCGCTCCGATACGAGACCGCGCC 59

RESULT 12

US-10-424-599-89670/c
; Sequence 89670, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 89670
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_51983C.1
US-10-424-599-89670

Query Match 48.0%; Score 28.8; DB 16; Length 490;
Best Local Similarity 69.6%; Pred. No. 1.2;
Matches 39; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

OY 5 CAACAGCTCTCCCGCTCATGGGTTACTTCTGCTCCAGATACTGCCAGCTCCA 60
|||
Db 273 CAGCGCTTCTTGCGCTCACGCGCTTCTTGCGCTCCAGCGCTTGCTCGCTCCA 218

RESULT 13

US-10-335-394-42
; Sequence 42, Application US/10335394
; Publication No. US20030138452A1
; GENERAL INFORMATION:
; APPLICANT: Kwak, Larry
; APPLICANT: Biragyn, Arya
; TITLE OF INVENTION: METHODS AND COMPOSITIONS OF
; FILE REFERENCE: 14014.0316/P
; CURRENT APPLICATION NUMBER: US/10/335,394
; CURRENT FILING DATE: 2002-12-31
; PRIOR APPLICATION NUMBER: US/09/646,028
; PRIOR FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: 60/077,745
; PRIOR FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for windows Version 3.0
; SEQ ID NO 42
; LENGTH: 519
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of artificial sequence:/note=synthetic construct
US-10-335-394-42

Query Match 47.7%; Score 28.6; DB 15; Length 519;
Best Local Similarity 67.8%; Pred. No. 1.5;
Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

OY 1 GGGTCAACAGCTCTCCCGCTCATGGGTTACTTCTGCTCCAGATACTGCCAGCTCC 59
|||
Db 355 GGTTAACTTCTGCGCCCTCGACGGTTAACTTCTGCGCCGACACTCTGCCAGCACC 413

RESULT 14
US-10-335-394-46
; Sequence 46, Application US/10335394
; Publication No. US20030138452A1
; GENERAL INFORMATION:
; APPLICANT: Kwak, Larry
; APPLICANT: Biregyu, Arya
; TITLE OF INVENTION: METHODS AND COMPOSITIONS OF
; FILE REFERENCE: 14014.0316/P
; CURRENT APPLICATION NUMBER: US/10/335,394
; PRIOR FILING DATE: 2002-12-31
; PRIOR APPLICATION NUMBER: US/09/646,028
; PRIOR FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: 60/077,745
; PRIOR FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 46
; LENGTH: 534
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of artificial sequence:/note=synthetic construct
US-10-335-394-46

Query Match 47.7%; Score 28.6; DB 15; Length 534;
Best Local Similarity 67.8%; Pred. No. 1.5;
Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 GGGTCAACAGCTCCTCCGCTCATGGGTTACTTCTGCTCCAGATACGCCCCAGCTCC 59
DB 370 GGTGTAATTCTTGCCGCTCGACGGTGAATCTCTGCCCCGACACTGCCACGACC 428

RESULT 15
US-09-770-149-278/C
; Sequence 278, Application US/09770149
; Patent No. US20020059663A1
; GENERAL INFORMATION:
; APPLICANT: Goriach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Krickler, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2024 (PARA-013PRV)
; CURRENT APPLICATION NUMBER: US/09/770,149
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,506
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 278
; LENGTH: 697
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana

US-09-770-149-278

Query Match 46.3%; Score 27.8; DB 9; Length 697;
Best Local Similarity 69.1%; Pred. No. 3;
Matches 38; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 5 CAACAGCTCCTCCGCTCATGGGTTACTTCTGCTCCAGATACGCCCCAGCTCC 59
DB 373 CAAGTCTTCTGCGCTGAAGCGATGCTGCAGCTGCACTCGCAAGCTGC 319

Search completed: January 16, 2005, 09:30:24
Job time : 183.4 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 19:50:13 : Search time 35.2 Seconds
(without alignments)
1211.572 Million cell updates/sec

Title: US-10-057-136-6
Perfect score: 60
Sequence: 1 gggtcacacagctccctccgc.....cagatactcgccacgtcca 60

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 segs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/lna/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/lna/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/lna/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/lna/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/lna/PCtUS_COMB.seq: *
6: /cgn2_6/ptodata/1/lna/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	28.6	47.7	519	4	US-09-646-028-42 Sequence 42, Appl
2	28.6	47.7	534	4	US-09-646-028-46 Sequence 46, Appl
3	27.6	46.0	303	4	US-09-489-039A-2070 Sequence 2070, Ap
4	27	45.0	1804	4	US-10-029-517-17 Sequence 17, Appl
5	27	45.0	8186	4	US-10-029-517-19 Sequence 19, Appl
6	26.4	44.0	2301	1	US-08-306-691B-23 Sequence 23, Appl
7	26.4	44.0	2301	4	US-09-167-206-3 Sequence 3, Appl
8	26.4	44.0	2301	5	PCT-US93-06251-78 Sequence 78, Appl
9	26.4	44.0	16442	3	US-08-781-891-208 Sequence 208, App
10	26.4	44.0	16442	4	US-09-618-166-208 Sequence 208, App
11	26.2	43.7	1013	4	US-09-270-767-4588 Sequence 4588, Ap
12	26.2	43.7	1013	4	US-09-270-767-19870 Sequence 19870, A
13	25.6	42.7	205	4	US-09-513-999C-13924 Sequence 13924, A
14	25.6	42.7	453	4	US-09-221-017B-434 Sequence 434, App
15	25.6	42.7	518	4	US-10-029-517-101 Sequence 101, App
16	25.6	42.7	981	4	US-10-029-517-16 Sequence 16, Appl
17	25.6	42.7	1721	4	US-10-029-517-3 Sequence 3, Appl
18	25.6	42.7	1952	4	US-09-520-781-23 Sequence 23, Appl
19	25.6	42.7	2010	4	US-09-520-781-21 Sequence 21, Appl
20	25.4	42.3	572	4	US-10-029-517-18 Sequence 18, Appl
21	25.4	42.3	572	4	US-10-029-517-18 Sequence 18, Appl
22	25.4	42.3	1770	4	US-09-489-039A-3639 Sequence 3639, Ap
23	25.4	42.3	1230025	4	US-09-198-452A-1 Sequence 25147, A
24	25.2	42.0	418	4	US-09-270-767-25147 Sequence 25147, A
25	25.2	42.0	486	4	US-09-270-767-9852 Sequence 9852, Ap
26	25	41.7	2304	4	US-09-799-451-223 Sequence 223, App
27	25	41.7	6972	4	US-09-595-684B-38 Sequence 38, Appl

C	28	25	41.7	8309	4	US-09-620-312D-1083	Sequence 1083, Ap
	29	24.8	41.3	994	6	5202236-39	Patent No. 5202236
	30	24.8	41.3	1236	6	5202236-4	Patent No. 5202236
	31	24.8	41.3	1607	3	US-08-753-007A-3	Sequence 3, Appl
	32	24.8	41.3	1607	3	US-09-398-496-3	Sequence 3, Appl
	33	24.8	41.3	1926	4	US-09-249-585A-2	Sequence 2, Appl
	34	24.8	41.3	1926	4	US-09-410-399-3	Sequence 3, Appl
	35	24.8	41.3	2467	3	US-08-753-007A-1	Sequence 1, Appl
	36	24.8	41.3	2467	3	US-09-398-496-1	Sequence 1, Appl
	37	24.8	41.3	2580	3	US-09-050-863-2	Sequence 2, Appl
	38	24.8	41.3	2580	3	US-09-359-081-2	Sequence 1, Appl
	39	24.8	41.3	5452	2	US-09-130-114-1	Sequence 1, Appl
	40	24.8	41.3	8705	4	US-09-647-344A-14	Sequence 14, Appl
	41	24.8	41.3	9600	3	US-08-910-647-1	Sequence 1, Appl
	42	24.8	41.3	9600	4	US-09-620-925-1	Sequence 1, Appl
	43	24.8	41.3	10596	1	US-07-884-811-15	Sequence 15, Appl
	44	24.8	41.3	10596	1	US-07-885-971-15	Sequence 15, Appl
	45	24.8	41.3	10596	1	US-08-087-783A-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-09-646-028-42
Sequence 42, Application US/09646028
Patent No. 6562347
GENERAL INFORMATION:
APPLICANT: Kwak, Larry
APPLICANT: Biragyn, Arya
TITLE OF INVENTION: METHODS AND COMPOSITIONS OF
FILE REFERENCE: 14014.0316/P
CURRENT APPLICATION NUMBER: US/09/646, 028
CURRENT FILING DATE: 2000-09-12
PRIOR APPLICATION NUMBER: 60/077, 745
PRIOR FILING DATE: 1998-03-12
NUMBER OF SEQ ID NOS: 57
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 42
LENGTH: 519
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of artificial sequence:/note=synthetic construct
US-09-646-028-42

Query Match 47.7%; Score 28.6; DB 4; Length 519;
Best Local Similarity 67.8%; Pred. No. 0.67;
Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 GGGTCACAGCTCTCCCGCTCATGGGTTACTTCTGCTCCAGATACTGCCAGCTCC 59
Db 355 GGTTAACTTCTGCGCCCTCGACGGTGTAACTTCTGCGCCGACACTCGCCAGCACC 413

RESULT 2
US-09-646-028-46
Sequence 46, Application US/09646028
Patent No. 6562347
GENERAL INFORMATION:
APPLICANT: Kwak, Larry
APPLICANT: Biragyn, Arya
TITLE OF INVENTION: METHODS AND COMPOSITIONS OF
FILE REFERENCE: 14014.0316/P
CURRENT APPLICATION NUMBER: US/09/646, 028
CURRENT FILING DATE: 2000-09-12
PRIOR APPLICATION NUMBER: 60/077, 745
PRIOR FILING DATE: 1998-03-12
NUMBER OF SEQ ID NOS: 57
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 46

```

; LENGTH: 534
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of artificial sequence: /note=synthetic construct
US-09-646-028-46

```

Query Match	47.7%;	Score 28.6;	DB 4;	Length 534;
Best Local Similarity	67.8%;	Pred. No. 0.68;		
Matches	40;	Conservative	0;	Mismatches 19;
				Indels 0;
				Cars 0

QY 1 GGGTACAAGCTCCTCCCGCTCATGGGTTACTTGTGCCAGATACTCGCCCAAGTCC 59
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 370 GGTGTAACTTCTGC CGCCCTCGACGGTGTA ACTTCTGCCCCGGA CACTCGCCCAAGACC 428

RESULT 3
US-09-489-039A-2070/C
; Sequence 2070, Application US/09489039A
; Patent No. 6610836

```

; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 2070
; LENGTH: 303
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-2070

```

Query Match	46.0%;	Score 27.6;	DB 4;	Length 303;
Best Local Similarity	67.2%;	Pred. No. 1.3;		
Matches	39;	Conservative	0;	Mismatches 19.
				Indels

	conservative	0;	Mismatches	19;	Indels	0;	Gaps	0;
QY	1	GGGTCACAGCTCCTCCCGCTCATGGGGTTACTTGTGCATCCAGATACTGCCCAAGTC	58					
Dd	157	GGGTCAGCTGCCGACTGCAGCGCCTGGCGGATTTCCTTCACATCCTCACCCAGATC	100					

RESULT 4
US-10-029-517-17
; Sequence 17, Application US/10029517
Patent No. 7,515,500

```

; GENERAL INFORMATION:
; PATENT NO.: 5/1562/
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 17
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (73) ... (1500)
US-10-029-517-17

```

```
S-10-029-517-17
Query Match      45.0%; Score 27; DB 4; Length 1804;
Best Local Similarity 66.1%; Pred. No. 3.2;
Matches 39; Conservative 0; Mismatches 20; Indels 0; Gaps 0
```

	Conservative	Mismatches	Indels	Gaps
Oy	1	0	20	0
Db	457	0	0	515

RESULT 5
US-10-029-517-19
; Sequence 19, Application US/10029517
; Patent No. 6716627

; APPLICANT: Kenneth W. Dobie
 ; APPLICANT: Susan J. Myers
 ; TITLE OF INVENTION: ANTIGENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
 ; FILE REFERENCE: RTS-0352
 ;

```

; CURRENT FILING DATE: 2001-12-20
; CURRENT APPLICATION NUMBER: US/10/029,517
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 19

```

LENGTH: 8186

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```

```

; NAME/KEY: unsure
; LOCATION: 6899

```

```

; OTHER INFORMATION: unknown
; NAME/KEY: unsure

```

```

; LOCATION: 7155
; OTHER INFORMATION: unknown

```

```

;
; NAME/KEY: unsure
; LOCATION: 7184
;

```

```

; OTHER INFORMATION: unknown
; NAME/KEY: unsure
;

```

```

; LOCATION: 7957
; OTHER INFORMATION: unknown
;

```

```
; NAME/KEY: Intron
; LOCATION: (2997) ... (3498)
```

```

; OTHER INFORMATION: intron 1
;
; NAME/KEY: intron:exon junction
;
; LOCATOR: (3468) (3475)

```

```

LOCATION: (3498) ... (3499)
OTHER INFORMATION: intron 1:exon 2
NAME/KEY: 3498

```

NAME/NEI: EXOH
LOCATION: (3508) ... (3599)
OTHER INFORMATION: ...

NAME/KEY: exon:intron junction
LOCATION: (3983) (3993)

OTHER INFORMATION: exon 2a:intron 2a
NAME/KEY: intron:exon junction

LOCATION: (4205) .. (4206)
OTHER INFORMATION: intron 3d:xxxx 3a

NAME/KEY: Intron:exon junction
LOCATION: (4259) (4260)

OTHER INFORMATION: intron 2d:exon 3d
NAME/KEY: exon

LOCATION: (4260) ... (4328)
OTHER INFORMATION: exon 3d

NAME/KEY: intron:exon junction
LOCATION: (4632)..(4633)

OTHER INFORMATION: intron 3:exon 4
NAME/KEY: exon

LOCATION: (4914) ... (5035)
OTHER INFORMATION: exon 5

NAME/KEY:	INTRON
LOCATION:	(5266) ... (6293)

OTHER INFORMATION: intron 6
S-10-029-517-19

Query Match	45.0%;	Score 27;	DB 4;	Length 8186;
Best Local Similarity	66.1%;	Pred. No. 4.6;		
Matches 39;	Conservative 0;	Mismatches 20;	Indels 0;	Gaps 0.

OY 1 GGGTCAACAGCTCCTCCCGCTCATGGGTTACTTGTGTCAGATACTGCCCGCAGCTCC 59
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3825 GGCTTCACCGCCCCCCCAGCGCCACGGTGTCACCTCGGCCCCGGAACACGAGCGCGCCCC 3883

RESULT 6


```

US-08-306-691B-23/c
; Sequence 23, Application US/08306691B
; Patent No. 5734039
; GENERAL INFORMATION:
; APPLICANT: Calabretta, Bruno
; APPLICANT: Skorski, Tomasz
; TITLE OF INVENTION: ANTISENSE
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavoigna & Monaco, P.C.
; STREET: Two Penn Center, Suite 1800
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 KB
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/306,691B
; FILING DATE: September 15, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8321-8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5734039e
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2301 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-08-306-691B-23

Query Match          44.0%; Score 26.4; DB 1; Length 2301;
Best Local Similarity 69.2%; Pred. No. 5.5;
Matches 36; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Cy      8 CAGCTCCTCCCGCTCATGGGTTACTTCTGTCCAGATACTCGCCAGCTCC 59
        ||||| | | | | | | | | | | | | | | | | | | | | | |
Db     85 CTGCTCCTCCTGCTCCTCTCTGCTCTCTGCTCTCTGCTCTCTGCTCTGCTCC 34

RESULT 7
; Sequence 3, Application US/09167206A
; Patent No. 6476193
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Schulz, Vincent P.
; APPLICANT: Yang, Meijia
; TITLE OF INVENTION: NIK1 PROTEIN AND NIK1 PROTEIN COMPLEXES
; FILE REFERENCE: 15966-521 Nik1 protein complexes
; CURRENT APPLICATION NUMBER: US/09/167,206A
; CURRENT FILING DATE: 1998-10-06
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patentin Ver. 2.0 .
; SEQ ID NO 3
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS

```

```

; LOCATION: (233) .. (2155)
US-09-167-206-3

Query Match          44.0%; Score 26.4; DB 4; Length 2301;
Best Local Similarity 69.2%; Pred. No. 5.5;
Matches 36; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY      8 CAGCTCCTCCCGCTCATGGGTTACTTCTGCTCCAGATACTGCCAGCTCC 59
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      85 CTGCTCCTCTGCTCTCCTCTCCTGCTCTCCTGCTCTCCTGCTCTCCTGCTCC 34

RESULT 8
PCT-US93-06251-78/C
; Sequence 78, Application PC/TUS9306251
; GENERAL INFORMATION:
; APPLICANT: Wickstrom, Eric and Rife, Jason P.
; TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
; TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: NY
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06251
; FILING DATE: 19930630
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8586
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 78:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2301 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-78

Query Match          44.0%; Score 26.4; DB 5; Length 2301;
Best Local Similarity 69.2%; Pred. No. 5.5;
Matches 36; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY      8 CAGCTCCTCCCGCTCATGGGTTACTTCTGCTCCAGATACTGCCAGCTCC 59
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      85 CTGCTCCTCTGCTCTCCTCTCCTGCTCTCCTGCTCTCCTGCTCTCCTGCTCC 34

RESULT 9
US-08-781-891-208
; Sequence 208, Application US/08781891
; Patent No. 6090620
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Yu, Chang-En
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; APPLICANT: Schellenberg, Gerald D.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO

```

```

1  TITLE OF INVENTION:  WERNER'S SYNDROME
2  NUMBER OF SEQUENCES:  209
3  CORRESPONDENCE ADDRESS:
4  ADDRESSEE:  SEED and BERRY LLP
5  STREET:  6300 Columbia Center, 701 Fifth Avenue
6  CITY:  Seattle
7  STATE:  Washington
8  COUNTRY:  USA
9  ZIP:  98104-7092
10 COMPUTER READABLE FORM:
11 MEDIUM TYPE:  Floppy disk
12 COMPUTER:  IBM PC compatible
13 OPERATING SYSTEM:  PC-DOS/MS-DOS
14 SOFTWARE:  PatentIn Release #1.0, Version #1.30
15 CURRENT APPLICATION DATA:
16 APPLICATION NUMBER:  US/08/781,891
17 FILING DATE:  27-DEC-1996
18 CLASSIFICATION:  800
19 ATTORNEY/AGENT INFORMATION:
20 NAME:  No. 6090620tenburg Ph.D., Carol
21 REGISTRATION NUMBER:  39,317
22 REFERENCE/DOCKET NUMBER:  240052.419
23 TELECOMMUNICATION INFORMATION:
24 TELEPHONE:  (206) 622-4900
25 TELEFAX:  (206) 682-6031
26 INFORMATION FOR SEQ ID NO:  208:
27 SEQUENCE CHARACTERISTICS:
28 LENGTH:  16442 base pairs
29 TYPE:  nucleic acid
30 STRANDEDNESS:  single
31 TOPOLOGY:  linear
32
33 US-08-781-891-208

```

[illegible]

RESULT 10
US-09-618-166-208
; Sequence 208, Application US/09618166
; Patent No. 658312
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; Yu, Chang-Bn
; Oshima, Junko
; Mulligan, John T.
; Schellenberg, Gerald D.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed Intellectual Property Law Group
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/618,166
; FILING DATE: 17-Jul-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: McMasters, David D.

```

;      REGISTRATION NUMBER: 33,963
;      REFERENCE/DOCKET NUMBER: 240052.419C1
;      TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (206) 622-4900
;      TELEFAX: (206) 682-6031
;      INFORMATION FOR SEQ ID NO: 208:
;
;      SEQUENCE CHARACTERISTICS:
;      LENGTH: 16442 base pairs
;      TYPE: nucleic acid
;      STRANDEDNESS: single
;      TOPOLOGY: linear
;      SEQUENCE DESCRIPTION: SEQ ID NO: 208:
US-09-618-166-208

```

Query Match	44.0%;	Score 26.4;	DB 4;	Length 16442;
Best Local Similarity	69.2%;	Pred. No. 8.7;		
Matches 36; Conservative	0;	Mismatches 16;	Indels 0;	Gaps 0;

QY 8 CAGCTCCTCCCGCTCATGGGTTACTTCTGCTCCAGATACTCGGCCACGTCC 59
 | | | | | | | | | | | | | | | | | | | |
Db 16316 CTGCTCTCTCTCTCTCTCTGTCCTCTCTGCTCCTGCTCTCTCTCTCTGCTCC 16367

```

RESULT 11
US-09-270-767-4588
; Sequence 4588, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4588
; LENGTH: 1013
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-4588

```

Query Match	43.7%	Score 26.2;	DB 4;	Length 1013;
Best Local Similarity	67.3%	Pred. No. 5.3;		
Matches 37; Conservative	0;	Mismatches 18;	Indels 0;	Gaps 0;

QY 5 CAACAGCTCCTCCGCTCATGGGGTTACTTGTGCTCCAGATACTCGCCAGCTCC 59
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13 CAGCTTCAGCTCCAGCTCCTGGGGTTCTTCTGTACCAGAAGTCCGTGCCAGTACC 67

```

RESULT 12
US-09-270-767-19870
; Sequence 19870, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19870
; LENGTH: 1013
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-19870

```

Query Match	43.7%	Score 26.2;	DB 4;	Length 1013;
Best Local Similarity	67.3%	Pred. No. 5.3;		
Matches 37; Conservative	0;	Mismatches 18;	Indels 0;	Gaps 0;

5 CAACAGCTCCTCCCGCTCATGGGTTACTTCTGCTCCAGATACTGCCCAAGCTCC 59

Db 13 CAGCTCAGCTCCAGCTCCTGGGTTCTTGTACCAAGTCTGTTCCAGTACC 67

RESULT 13

US-09-513-999C-13924/c
; Sequence 13924, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513, 999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 13924
; LENGTH: 205
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-13924

Query Match 42.7%; Score 25.6; DB 4; Length 205;
Best Local Similarity 66.1%; Pred. No. 5.8;
Matches 37; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 5 CAACAGCTCCTCCCGCTCATGGGTTACTTGTGCTCCAGATACTCGCCAGCTCCA 60
Db 138 CAGCGCTTCTGCGCTCCAGCGGCTTCTGCGCTCCAGCGGCTTCTGCGCTCCA 83

RESULT 14

US-09-221-017B-434/c
; Sequence 434, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1546
; FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP2911
; FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU98/01023
; FILING DATE: 10-DEC-1998
; ATTORNEY/AGENT INFORMATION:

NAME: Monroy, Gladys H
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-20021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 434:
SEQUENCE CHARACTERISTICS:
LENGTH: 453 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORYPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1...453
US-09-221-017B-434

Query Match 42.7%; Score 25.6; DB 4; Length 453;
Best Local Similarity 70.8%; Pred. No. 7;
Matches 34; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 9 AGCTCTCCCGCTCATGGGTTACTTGTGCTCCAGATACTCGCCAGC 56
Db 420 AGCTTCTTCCGCTCAGCGGGGTATATCCGTGCGGCTGTGCCCCATC 373

RESULT 15

US-10-029-517-101
; Sequence 101, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 101
; LENGTH: 518
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-10-029-517-101

Query Match 42.7%; Score 25.6; DB 4; Length 518;
Best Local Similarity 66.1%; Pred. No. 7.2;
Matches 37; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 GGTCAACAGCTCCTCCCGCTCATGGGTTACTTGTGCTCCAGATACTCGCCAGC 56
Db 462 GGTCAACAGCTCCTCCCGCTCATGGGTTACTTGTGCTCCAGATACTCGCCAGC 517

Search completed: January 16, 2005, 03:01:28
Job time : 37.2 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 19:43:55 ; Search time 1550.6 Seconds
(without alignments)
1410.024 Million cell updates/sec

Title: US-10-057-136-6

Perfect score: 60

Sequence: 1 GGGTCACAGCTCTCTCCCGC.....CAGATACCTGCGCCAGCTCCA 60

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	32.6	54.3	603	CF809795	CF809795 Lr LC1ED
2	32	53.3	418	AZ943545	AZ943545 2M0204E18
3	30.4	50.7	652	CNS03USR	AL261540 Tetraodon
C 4	30.4	50.7	983	CNS03W76	AL263355 Tetraodon
C 5	29.8	49.7	1101	CNS057K8	AL324737 Tetraodon
6	29.6	49.3	413	BJ511928	BJ511928 BJ511928
C 7	29.6	49.3	568	BJ498303	BJ498303 BJ498303
C 8	29.6	49.3	584	BJ500902	BJ500902 BJ500902
C 9	29.6	49.3	612	BJ493595	BJ493595 BJ493595
C 10	29.6	49.3	652	BJ014441	BJ014441 BJ014441
C 11	29.6	49.3	652	BJ490836	BJ490836 BJ490836
C 12	29.6	49.3	669	BJ504955	BJ504955 BJ504955
13	29.6	49.3	686	BJ028611	BJ028611 BJ028611
C 14	29.6	49.3	697	BJ710454	BJ710454 BJ710454
15	29.6	49.3	730	BJ508359	BJ508359 BJ508359
16	29.6	49.3	739	BJ516738	BJ516738 BJ516738
17	29.6	49.3	798	BJ721686	BJ721686 BJ721686
18	29.4	49.0	644	CNS0ABAA	BX818432 Arabidops
19	29.2	48.7	684	CA094017	CA094017 SCCCL400
20	29.2	48.7	684	CA083318	CA083318 SCEPAM201
21	29.2	48.7	731	CA289703	CA289703 SCAGFL800
22	29.2	48.7	871	CL566510	CL566510 OB_Ba003
23	29	48.3	209	CE134399	CE134399 tigr-gss-
C 24	29	48.3	341	AA591111	AA591111 vm12c06.r

C 25	29	48.3	345	1	AA967806	AA967806 uh05d06.r
C 26	29	48.3	380	7	W65797	W65797 me14g02.r1
C 27	29	48.3	393	1	AA153891	AA153891 mq56e05.r
C 28	29	48.3	436	2	BB863659	BB863659 BB863659
C 29	29	48.3	454	1	AI892918	AI892918 mq56e05.y
C 30	29	48.3	523	5	BQ745334	BQ745334 UI-M-ER0-
C 31	29	48.3	838	4	CB201841	CB201841 AGENCOURT
C 32	29	48.3	874	6	BI525902	BI525902 602924866
C 33	29	48.3	893	5	BQ939292	BQ939292 AGENCOURT
C 34	29	48.3	1295	9	AY411302	AY411302 Mus muscu
35	28.8	48.0	684	9	CNS04LNB	AL296333 Tetraodon
36	28.8	48.0	789	9	CNS02HOS	AL197029 Tetraodon
C 37	28.8	48.0	809	9	CNS01VGB	AL168716 Tetraodon
C 38	28.8	48.0	844	9	AG551682	AG551682 Mus muscu
C 39	28.6	47.7	330	1	AI925867	AI925867 wo20d04.x
40	28.6	47.7	430	5	BY258590	BY258590 BY258590
41	28.6	47.7	471	7	H19186	H19186 ym50g03.r1
C 42	28.6	47.7	644	6	CA370120	CA370120 650203 NC
43	28.6	47.7	1113	5	BU148487	BU148487 AGENCOURT
44	28.6	47.7	1234	5	BQ936898	BQ936898 AGENCOURT
45	28.6	47.7	1262	5	BQ935496	BQ935496 AGENCOURT

ALIGNMENTS

RESULT 1
CF809795/c
LOCUS
DEFINITION Lr LC1ED_31G07 SAC Lumbricus rubellus late Cocoon Library 1
Lumbricus rubellus CDNA clone Lr LC1ED_31G07 5', mRNA sequence.
CF809795
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

EST.
Lumbricus rubellus (humus earthworm)
Lumbricus rubellus
Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haptotaxida;
Lumbricina; Lumbricidae; Lumbricus.

REFERENCE
1 (bases 1 to 603)
Jones, M., Chaselay, J., Hedley, B.A., Morgan, J.C., Sturzenbaum, S.,
Kille, P. and Blaxter, M.

TITLE
The Lumbricus rubellus EST program - Sequences from a cocoon
library

JOURNAL
COMMENT
Unpublished (2003)
Contact: Blaxter ML

Institute of Cell, Animal and Population Biology
University of Edinburgh
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9

3JT, UK.
Tel: +44 131 650 6760
Fax: +44 131 670 5450

Email: mark.blaxter@ed.ac.uk
The library was prepared using protocols given by the supplier
(Stratagene).

PCR Primers
FORWARD: M13R(ACCGATAACAATTTCACACAGA)
BACKWARD: T7PL(CTCACTATAGGCGAATTGG)
Plate: 31 row: G column: 07

Seq primer: SAC(GGGAACAAGCTGAG)
High quality sequence stop: 504.

FEATURES

source

Location/Qualifiers
1..603
/organism="Lumbricus rubellus"
/mol_type="mRNA"
/db_xref="taxon:35632"
/clone="Lr LC1ED_31G07"
/tissue_type="whole worm"
/dev_stage="Late Cocoon"
/clone_lib="Lumbricus rubellus late Cocoon Library 1"
/note="Vector: pBluescript II SK+; The library was
prepared using protocols given by the supplier
(Stratagene)."

ORIGIN

Query Match 54.3%; Score 32.6; DB 7; Length 603;
Best Local Similarity 74.5%; Pred. No. 4.5;
Matches 41; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 4 TCACAGCTCCTCCCGCTCATGGGTTACTTCTGTCTCCAGATACTCGCCAGCTC 58
|||||
Db 566 TCAGCAGCTCCTCCTCTCTCACTCGGTGCTCCTGCTCTCTCTCTCCACCTC 512
|||||

RESULT 2
AZ943545
LOCUS

DEFINITION 418 bp DNA linear GSS 26-APR-2001
2M0204E18F Mouse 10kb plasmid UUGC2M library Mus musculus genomic
clone UUGC2M0204E18 F, genomic survey sequence.

ACCESSION AZ943545
VERSION AZ943545.1 GI:13807673
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 418)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
Niederhausen,A. and Wright,D.,Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA

Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0204 row: E column: 18
Seq primer: CGTTGTAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 418.

FEATURES
source location/Qualifiers
1..418

/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0204E18"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10Kb plasmid UUGC2M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (female) was obtained from the Jackson
laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (gi|4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN

Query Match 53.3%; Score 32; DB 8; Length 418;
Best Local Similarity 73.2%; Pred. No. 6.8;
Matches 41; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 5 CAACAGCTCCTCCCGCTCATGGGTTACTTCTGTCTCCAGATACTCGCCAGCTCCA 60
|||||
Db 53 CAGCTCCTCCTCCTGCTCTGCTCTCTCTCTCCAGCTCCAGCTACTCTCTCCAGCTCCA 108
|||||

RESULT 3
CNS03USR
LOCUS

DEFINITION 652 bp DNA linear GSS 01-SEP-2000
Tetraodon nigroviridis genome survey sequence T7 end of clone
059N19 of library G from Tetraodon nigroviridis, genomic survey
sequence.

ACCESSION AL261540
VERSION AL261540.1 GI:7983165
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis

REFERENCE 1
AUTHORS Roest Crolius,H., Jalllon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.

TITLE

JOURNAL Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)

MEDLINE 20296633
PUBMED 10835645

TITLE

JOURNAL Roest Crolius,H., Jalllon,O., Dasilva,C., Ozouf-Costez,C.,
Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
Saurin,W., Bernot,A. and Weissenbach,J.

MEDLINE 20359837
PUBMED 10899143
REFERENCE 3 (bases 1 to 652)
Genoscope.
Genoscope.

REFERENCE Direct Submission
AUTHORS Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT

This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.

FEATURES
source location/Qualifiers
1..652

/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="059N19"
/clone_lib="G"
/note="Genoscope sequence ID : COBG059CG10LP1-end : T7"

ORIGIN

Query Match 50.7%; Score 30.4; DB 9; Length 652;
Best Local Similarity 71.4%; Pred. No. 23;
Matches 40; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 5 CAACAGCTCCTCCCGCTCATGGGTTACTTCTGTCTCCAGATACTCGCCAGCTCCA 60
|||||
Db 136 CACCTGCTCCTCTGCTCA 191
|||||

RESULT 4

CNS03W76/c
LOCUS CNS03W76 983 bp DNA linear GSS 01-SEP-2000
DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone
063L11 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL263355 GI:7985013
VERSION AL263355
KEYWORDS GSS: genome survey sequence.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
REFERENCE 1
AUTHORS Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
TITLE Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
JOURNAL Nat. Genet. 25 (2), 235-238 (2000)
MEDLINE 20296633
PUBMED 10835645
REFERENCE 2
AUTHORS Roest Crolius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
Saurin,W., Bernot,A. and Weissenbach,J.
TITLE Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
JOURNAL Genome Res. 10 (7), 939-949 (2000)
MEDLINE 20359837
PUBMED 10899143
REFERENCE 3 (bases 1 to 983)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
FEATURES
source
1..983
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="063L11"
/clone_1lb="G"
/note="Genoscope sequence ID : C0BG063CF06LP1-end : T7"
ORIGIN
Query Match 50.7%; Score 30.4; DB 9; Length 983;
Best Local Similarity 71.4%; Pred. No. 25;
Matches 40; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 5 CAACAGCTCTCCCGCTCATGGGTTACTTGTGCTCCAGATACTCGCCAGCTCCA 60
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 862 CACCTGCTCA 807
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 5
CNS057K8/c 1101 bp DNA linear GSS 26-JUL-2000
LOCUS CNS057K8/c
DEFINITION Tetraodon nigroviridis genome survey sequence SP6 end of clone
002K02 of library B from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL324737 GI:9557619
VERSION AL324737
KEYWORDS GSS: genome survey sequence.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
REFERENCE 1
AUTHORS Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
TITLE Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
JOURNAL Nat. Genet. 25 (2), 235-238 (2000)
MEDLINE 20296633
PUBMED 10835645
REFERENCE 2
AUTHORS Roest Crolius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
Saurin,W., Bernot,A. and Weissenbach,J.
TITLE Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
JOURNAL Genome Res. 10 (7), 939-949 (2000)
MEDLINE 20359837
PUBMED 10899143
REFERENCE 3 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
FEATURES
source
1..1101
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="002K02"
/clone_1lb="B"
/note="Genoscope sequence ID : C0AB002BF01B1-end : SP6"
ORIGIN
Query Match 49.7%; Score 29.8; DB 9; Length 1101;
Best Local Similarity 70.2%; Pred. No. 39;
Matches 40; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 3 GTCAACAGCTCTCCCGCTCATGGGTTACTTGTGCTCCAGATACTCGCCAGCTCC 59
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 299 GACTTCACCTTTC 243
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 6
BJS11928 413 bp mRNA linear EST 08-AUG-2002
LOCUS BJS11928 MF01FSA cDNA Oryzias latipes cDNA clone MF01FSA042E01 3',
DEFINITION mRNA sequence.
ACCESSION BJS11928
VERSION BJS11928 GI:22163890
KEYWORDS EST.
SOURCE Oryzias latipes (Japanese medaka)
ORGANISM Oryzias latipes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
REFERENCE 1 (bases 1 to 413)
AUTHORS Kohara,Y., Shin-I,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
TITLE Medaka EST Project in Takeda's lab
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasu Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856

FEATURES
source 1. .413
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp
Location/Qualifiers

```

/organism="Oryzias latipes"
/mol_type="mRNA"
/strain="d-rR"
/db_xref="taxon:8090"
/clone="MF01FSA042E01"
/sex="mixture of female and male"
/tissue_type="whole embryo"
/dev_stage="fcy stage 40"
/clone_lib="MF01FSA cDNA"

```

Query Match	49.3%;	Score 29.6;	DB 4;	Length 413;
Best Local Similarity	68.3%;	Pred. No. 40;		
Matches 41; Conservative	0;	Mismatches 19;	Indels 0;	Gaps 0.

DY 1 GGGTCAACAGCTCCTCCCGGCTCATGGGGTTACTTGTCGCCAATACCTGCCAGCTCCA 60
| | | | | | | | | | | | | | | | | | | | | |
Db 309 GCGTCATGGGCTCCCGCTCTCATCGGATCTTCGGGTCCAGGTCGTGCACGAGCTTCCA 368

RESULT 7	LOCUS	DEFINITION	568 bp	mRNA	linear	EST 08-AUG-2002
	BJ498303/c					
	BJ498303	MF01FSA	CDNA	Oryzias latipes	CDNA	clone MF01FSA012104 5',
		mRNA	sequence.			

ACCESSION	BJ498303
VERSION	BJ498303.1
KEYWORDS	GI:22150265
SOURCE	EST.
ORGANISM	Oryzias latipes (Japanese medaka)
	Oryzias latipes

REFERENCE
1 (pages 1 to 569)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha; Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	(bases 1 to 568)	Kohara, Y., Shin-i, T., Kimura, T., Narita, T., Jindo, T. and Takeda, H.	Medaka EST Project in Takeda's lab	Unpublished (2001)
		Contact: Tadasu Shin-i		

Center for Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855

FEATURES
source
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1. .568

```

/organism="Oryzias latipes"
/mol_type="mRNA"
/strain="d-rR"
/db_xref="taxon:8090"
/clone="MF01FSA012104"
/sex="mixture of female and male"
/tissue_type="whole embryo"
/dev_stage="fry stage 40"
/clone_lib="MF01FSA cDNA"

```

ORIGIN	Score	DB	Length	Indels	Cons
Query Match	49.3%	29.6;	4;	568;	
Best Local Similarity	68.3%	Pred. No. 42;			
Matches 41;	Conservative	0;	Mismatches 19;	Indels 0;	Cons 0;

1 GGGTCACAGAGCTCTCCCGCATGCGGTACTTCTGCTCCAGATACGCCAGCTCCA 60
497 GCGTCATGGGCTCCCGCTCATCGGATCTTCGGGTCACAGGTCGTGACGAGCTCCA 438

RESULT 8
3J500902/c

LOCUS BJ500902 584 bp mRNA linear EST 08-AUG-2002
DEFINITION BJ500902 MF01FSA cDNA *Oryzias latipes* CDNA clone MF01FSA047C16 5',
mRNA sequence.
ACCESSION BJ500902
VERSION BJ500902.1 GI:22152864
KEYWORDS EST.
SOURCE *Oryzias latipes* (Japanese medaka)
ORGANISM *Oryzias latipes*
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
Belontiiformes; Adrianichthyidae; Oryziinae; *Oryzias*.
1 (bases 1 to 584)
Kohara, Y., Shin-i, T., Kimura, T., Narita, T., Jindo, T. and Takeda, H.
Medaka EST Project in Takeda's lab
Unpublished (2001)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
location/Qualifiers
1..584

```

/organism="Oryzias latipes"
/mol_type="mRNA"
/strain="d-rR"
/db_xref="taxon:8090"
/clone="MF01FSA047C16"
/sex="mixture of female and male"
/tissue_type="whole embryo"
/dev_stage="fry stage 40"
/clone_lib="MF01FSA cDNA"

```

Query Match	49.3%;	Score 29.6;	DB 4;	Length 584;
Best Local Similarity	68.3%;	Pred. No. 42;		
Matches 41; Conservative	0;	Mismatches 19;	Indels 0;	Gaps 0.

Dy 1 GGGTCAACAGCTCCTCCGCTCATGGGGTTACTTGCTCCAGATACTCGCCAGCTCCA 60
| | | | | | | | | | | | | | | | | | | | |
Db 524 GCGTCATGGGCTCCCCGCTCTCATCCGGATCTTCGGGCTCCAGCTGTCTCAGAGCTCCA 465

[illegible]

ACCESSION	BJ493595	
VERSION	BJ493595.1	GI:22145521
KEYWORDS	EST.	
SOURCE	Oryzias latipes	(Japanese medaka)
ORGANISM	Oryzias latipes	

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha; Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.

REFERENCE 1 (bases 1 to 612)
AUTHORS Kohara, Y., Shin-i, T., Kimura, T., Narita, T., Jindo, T. and Takeda, H.
TITLE Medaka EST Project in Takeda's lab
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasu Shin-i

Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856
Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp

FEATURES	Location/Qualifiers
source	1. .612

/organism="Oryzias latipes"

/organism="Oryzias latipes"

```
/mol_type="mRNA"
/db_xref="taxon:8090"
/clone="MF01FSA032005"
/sex="mixture of female and male"
/tissue_type="whole embryo"
/dev_stage="fry stage 40"
/clone_1lb="MF01FSA cDNA"

ORIGIN

Query Match          49.3%; Score 29.6; DB 4; Length 612;
Best Local Similarity 68.3%; Pred. No. 42;
Matches 41; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 GGGTCACAGCTCTCTCCGCTCATGGGCTTACTTCTGCTCCAGATACTCGCCAGCTCCA 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 437 GCGTCATGGGCTCCCGCTCTCATCGGATCTTCGGGCTCCAGTCTGTCGACGAGCTCCA 378

RESULT 10
BJ014441/c          652 bp mRNA linear EST 05-DEC-2001
LOCUS               BJ014441 MF01SSA cDNA Oryzias latipes cDNA clone MF01SSA199H08 5',
DEFINITION          mRNA sequence.
ACCESSION           BJ014441
VERSION             BJ014441.1 GI:17373347
KEYWORDS            EST.
SOURCE              Oryzias latipes (Japanese medaka)
ORGANISM            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                    Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
                    Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
                    Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
REFERENCE            1 (bases 1 to 652)
AUTHORS             Kohara,Y., Shin-I,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
TITLE              Medaka EST Project in Takeda's lab
JOURNAL             Unpublished (2001)
COMMENT             Contact: Tadasu Shin-i
                    Center For Genetic Resource Information
                    National Institute of Genetics
                    1111 Yata, Mishima, Shizuoka 411-8540, Japan
                    Tel: 81-559-81-6856
                    Fax: 81-559-81-6855
                    Email: tshini@genes.nig.ac.jp.
                    Location/Qualifiers
FEATURES            1..652
source              /organism="Oryzias latipes"
                    /mol_type="mRNA"
                    /strain="Hd-rR"
                    /db_xref="taxon:8090"
                    /clone="MF01SSA199H08"
                    /sex="mixture of female and male"
                    /tissue_type="whole embryo"
                    /dev_stage="segmentation stage 20 - 25"
                    /clone_1lb="MF01SSA cDNA"

ORIGIN

Query Match          49.3%; Score 29.6; DB 4; Length 652;
Best Local Similarity 68.3%; Pred. No. 42;
Matches 41; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 GGGTCACAGCTCTCTCCGCTCATGGGCTTACTTCTGCTCCAGATACTCGCCAGCTCCA 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 517 GCGTCATGGGCTCCCGCTCTCATCGGATCTTCGGGCTCCAGTCTGTCGACGAGCTCCA 458

RESULT 11
BJ490836/c          652 bp mRNA linear EST 09-AUG-2002
LOCUS               BJ490836 MF01FSA cDNA Oryzias latipes cDNA clone MF01FSA021P19 5',
DEFINITION          mRNA sequence.
ACCESSION           BJ490836
VERSION             BJ490836.1 GI:22169585
```

```
KEYWORDS            EST.
SOURCE              Oryzias latipes (Japanese medaka)
ORGANISM            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                    Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
                    Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
                    Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
REFERENCE            1 (bases 1 to 652)
AUTHORS             Kohara,Y., Shin-I,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
TITLE              Medaka EST Project in Takeda's lab
JOURNAL             Unpublished (2001)
COMMENT             Contact: Tadasu Shin-i
                    Center For Genetic Resource Information
                    National Institute of Genetics
                    1111 Yata, Mishima, Shizuoka 411-8540, Japan
                    Tel: 81-559-81-6856
                    Fax: 81-559-81-6855
                    Email: tshini@genes.nig.ac.jp.
                    Location/Qualifiers
FEATURES            1..652
source              /organism="Oryzias latipes"
                    /mol_type="mRNA"
                    /strain="d-rR"
                    /db_xref="taxon:8090"
                    /clone="MF01FSA021P19"
                    /tissue_type="whole embryo"
                    /dev_stage="fry stage 40"
                    /clone_1lb="MF01FSA cDNA"

ORIGIN

Query Match          49.3%; Score 29.6; DB 4; Length 652;
Best Local Similarity 68.3%; Pred. No. 42;
Matches 41; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 GGGTCACAGCTCTCTCCGCTCATGGGCTTACTTCTGCTCCAGATACTCGCCAGCTCCA 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 526 GCGTCATGGGCTCCCGCTCTCATCGGATCTTCGGGCTCCAGTCTGTCGACGAGCTCCA 467

RESULT 12
BJ504955            669 bp mRNA linear EST 08-AUG-2002
LOCUS               BJ504955 MF01FSA cDNA Oryzias latipes cDNA clone MF01FSA012104 3',
DEFINITION          mRNA sequence.
ACCESSION           BJ504955
VERSION             BJ504955.1 GI:22156917
KEYWORDS            EST.
SOURCE              Oryzias latipes (Japanese medaka)
ORGANISM            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                    Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
                    Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
                    Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
REFERENCE            1 (bases 1 to 669)
AUTHORS             Kohara,Y., Shin-I,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
TITLE              Medaka EST Project in Takeda's lab
JOURNAL             Unpublished (2001)
COMMENT             Contact: Tadasu Shin-i
                    Center For Genetic Resource Information
                    National Institute of Genetics
                    1111 Yata, Mishima, Shizuoka 411-8540, Japan
                    Tel: 81-559-81-6856
                    Fax: 81-559-81-6855
                    Email: tshini@genes.nig.ac.jp.
                    Location/Qualifiers
FEATURES            1..669
source              /organism="Oryzias latipes"
                    /mol_type="mRNA"
                    /strain="d-rR"
                    /db_xref="taxon:8090"
                    /clone="MF01FSA012104"
                    /sex="mixture of female and male"
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 16:00:11 : Search time 170.5 Seconds
(without alignments)
1847.304 Million cell updates/sec

Title: US-10-057-136-6

Perfect score: 60

Sequence: 1 gggtcacagctcctcccgccgc.....cagatactcgcccgctcca 60

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_23Sep04:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	60	100.0	60	2	AAV48319	Aav48319 Nucleotid
2	38.2	63.7	156	10	ADK68635	Adk68635 HSP65-MUC
3	38.2	63.7	1800	10	ADK68629	Adk68629 HSP65-MUC
4	36.6	61.0	60	2	AAV48324	Aav48324 Nucleotid
5	36	60.0	60	2	AAV48323	Aav48323 Nucleotid
6	33.6	56.0	120	10	ADK68631	Adk68631 HSP65-MUC
7	31.8	53.0	60	2	AAV48318	Aav48318 Nucleotid
8	31.6	52.7	525	3	AAD00385	Aad00385 Human Muc
9	31.6	52.7	891	3	AAD00391	Aad00391 Ubiqutin
10	31.6	52.7	1371	3	AAD00388	Aad00388 Human Muc
11	31.6	52.7	1737	3	AAD00394	Aad00394 Ubiqutin
12	30.2	50.3	60	2	AAV48320	Aav48320 Nucleotid
13	30.2	50.3	68	4	AAS01385	Aas01385 SBMV coat
14	30	50.0	60	2	AAV48321	Aav48321 Nucleotid
15	30	50.0	60	2	AAV48322	Aav48322 Nucleotid
16	29.6	49.3	93	8	ABV77126	Abv77126 Oligonuci
17	29.6	49.3	157	8	ABV77127	Abv77127 Nucleotid
18	29.4	49.0	78	4	AAS01400	Aas01400 RCNMV coa
19	29	48.3	60	2	AAV48325	Aav48325 Nucleotid
20	29	48.3	78	4	AAS01403	Aas01403 RCNMV coa
21	28.8	48.0	68	4	AAS01387	Aas01387 SBMV coat

22	28.6	47.7	519	2	AAZ21162	Aaz21162 Artificia
23	28.6	47.7	534	2	AAZ21166	Aaz21166 Artificia
24	28.2	47.0	78	4	AAS01401	Aas01401 RCNMV coa
25	28.2	47.0	4144	12	AD157666	Ad157666 Human bre
26	28	46.7	68	4	AAS01386	Aas01386 SBMV coat
27	28	46.7	68	4	AAS01384	Aas01384 SBMV coat
28	28	46.7	78	4	AAS01399	Aas01399 RCNMV coa
29	28	46.7	78	4	AAS01402	Aas01402 RCNMV coa
30	28	46.7	83	4	AAS01395	Aas01395 LTVS coat
31	27.8	46.3	697	6	ABO65701	Abq65701 Arabidops
32	27.6	46.0	303	11	ACH96275	Ach96275 Klebsiell
33	27.6	46.0	2469	6	AAD41965	Aad41965 Murine vo
34	27.6	46.0	2469	10	ADE06927	Ade06927 Mouse L-t
35	27.4	45.7	60	2	AAV48317	Aav48317 Nucleotid
36	27.4	45.7	83	4	AAS01393	Aas01393 LTVS coat
37	27.4	45.7	435	9	ACH41656	Ach41656 Human foe
38	27.4	45.7	2160	12	ADH48430	Adh48430 Human KPP
39	27.4	45.7	2448	10	ADF82749	Adf82749 Leukaemia
40	27.2	45.3	78	4	AAS01398	Aas01398 RCNMV coa
41	27.2	45.3	281	4	ABL14771	Ab114771 Drosophil
42	27.2	45.3	2193	4	ABL08730	Ab108730 Drosophil
43	27.2	45.3	2281	4	ABL14770	Ab114770 Drosophil
44	27	45.0	48	4	AAS01380	Aas01380 Human muc
45	27	45.0	60	2	AAV48316	Aav48316 Nucleotid

ALIGNMENTS

RESULT 1	
AAV48319	
ID	AAV48319 standard; DNA; 60 BP.
XX	
AC	AAV48319;
XX	
DT	20-NOV-1998 (first entry)
XX	
DE	Nucleotide sequence encoding MUC1 tandem repeat unit R4.
XX	
KW	SB; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW	tumour; tumour-associated antigen.
XX	
OS	Homo sapiens.
XX	
PN	W09837095-A2.
XX	
PD	27-AUG-1998.
XX	
PF	24-FEB-1998; 98WO-US003693.
XX	
PR	24-FEB-1997; 97US-0038253P.
XX	
PA	(THER-) THERION BIOLOGICS CORP.
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.
PA	(DAND) DANA FARBER CANCER INST INC.
XX	
PI	Schlom J, Kantor J, Kuife D, Panicali D, Gritz L;
XX	
DR	WPI; 1998-467492/40.
XX	
PT	New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT	immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX	
PS	Disclosure; Page 11; 42pp; English.
XX	
CC	The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC	immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC	(RPV). The RPV was used in a pharmaceutical composition also containing
CC	an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC	recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC	does not undergo significant genetic deletion, thereby providing an
CC	unexpectedly stable and immunogenic pox virus. They can be used to
CC	prevent or treat tumours expressing MUC1 tumour-associated antigens

XX SQ Sequence 60 BP; 10 A; 23 C; 13 G; 14 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 60; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 3.7e-10;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GGGTCAACAGCTCCTCCGCTCATGGGTTACTTCTGCTCCAGATATCGCCAGCTCCA 60
11-SEP-2002. 1 GGGTCAACAGCTCCTCCGCTCATGGGTTACTTCTGCTCCAGATATCGCCAGCTCCA 60
Db 1 GGGTCAACAGCTCCTCCGCTCATGGGTTACTTCTGCTCCAGATATCGCCAGCTCCA 60
RESULT 2
ADK68635/c ADK68635 standard; cDNA; 156 BP.
XX
AC ADK68635;
XX
DT 06-MAY-2004 (first entry)
XX
DE HSP65-MUC1 antigen CTL epitope related cDNA #3.
XX
KW Human breast cancer; Mycobacterium bovis; heat shock protein 65; BCG;
KW HSP65; cytotoxic T lymphocyte; CTL; MUC1; HSP65-MUC1 antigen CTL epitope;
KW ss.
XX
OS Unidentified.
XX
PN CN1368384-A.
XX
PD 11-SEP-2002.
XX
PF 08-FEB-2001; 2001CN-00102614.
XX
PR 08-FEB-2001; 2001CN-00102614.
XX
PA (DIWE-) DIWEIHUAYU BIO TECHNOLOGY CO LTD BEIJING.
XX
PI Yu Y, Li H;
XX
DR WPI; 2003-854662/80.
XX
PT Genetically engineered vaccine of MUC-1 antigen for human breast cancer.
XX
PS Example 2; Page 4 (Disclosure); 14pp; Chinese.
XX
CC The invention relates to a method of preparation of a genetically
CC engineered vaccine for preventing and treating human breast cancer. The
CC method comprises fusing the coding gene of the Mycobacterium bovis heat
CC shock protein 65 (BCG HSP65) with the cytotoxic T lymphocyte (CTL)
CC epitope gene of MUC1 (antigen cell expressed by human breast cancer
CC cells) and the HSP65-MUC1 antigen CTL epitope is expressed in Escherichia
CC coli cells. This sequence represents DNA used in the method of the
CC invention.
XX
SQ Sequence 156 BP; 37 A; 42 C; 58 G; 19 T; 0 U; 0 Other;
OY Query Match 63.7%; Score 38.2; DB 10; Length 156;
Best Local Similarity 78.0%; Pred. No. 0.0059;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
Db 132 GGTCTACCGCTCCGCTCAGGGTGTACTCTGCTCCGACACCCGCTCCGCTCC 74
1 GGGTCAACAGCTCCTCCGCTCATGGGTTACTTCTGCTCCAGATATCGCCAGCTCC 59
132 GGTCTACCGCTCCGCTCAGGGTGTACTCTGCTCCGACACCCGCTCCGCTCC 74
RESULT 3
ADK68629 ADK68629 standard; cDNA; 1800 BP.
XX
AC ADK68629;
XX
DT 06-MAY-2004 (first entry)
XX
PA

XX DE HSP65-MUC1 antigen CTL epitope related cDNA #1.
XX
KW Human breast cancer; Mycobacterium bovis; heat shock protein 65; BCG;
KW HSP65; cytotoxic T lymphocyte; CTL; MUC1; HSP65-MUC1 antigen CTL epitope;
KW gene; ss.
XX
OS Unidentified.
XX
PN CN1368384-A.
XX
PD 11-SEP-2002.
XX
PF 08-FEB-2001; 2001CN-00102614.
XX
PR 08-FEB-2001; 2001CN-00102614.
XX
PA (DIWE-) DIWEIHUAYU BIO TECHNOLOGY CO LTD BEIJING.
XX
PI Yu Y, Li H;
XX
DR WPI; 2003-854662/80.
DR P-PSDB; ADK68630.
XX
PT Genetically engineered vaccine of MUC-1 antigen for human breast cancer.
XX
PS Claim 1; SEQ ID NO 1; 14pp; Chinese.
XX
CC The invention relates to a method of preparation of a genetically
CC engineered vaccine for preventing and treating human breast cancer. The
CC method comprises fusing the coding gene of the Mycobacterium bovis heat
CC shock protein 65 (BCG HSP65) with the cytotoxic T lymphocyte (CTL)
CC epitope gene of MUC1 (antigen cell expressed by human breast cancer
CC cells) and the HSP65-MUC1 antigen CTL epitope is expressed in Escherichia
CC coli cells. This sequence represents DNA used in the method of the
CC invention.
XX
SQ Sequence 1800 BP; 347 A; 561 C; 616 G; 276 T; 0 U; 0 Other;
OY Query Match 63.7%; Score 38.2; DB 10; Length 1800;
Best Local Similarity 78.0%; Pred. No. 0.0088;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
Db 1639 GGTCTACCGCTCCGCTCAGGGTGTACTCTGCTCCGACACCCGCTCCGCTCC 1697
1 GGGTCAACAGCTCCTCCGCTCATGGGTTACTTCTGCTCCAGATATCGCCAGCTCC 59
1639 GGTCTACCGCTCCGCTCAGGGTGTACTCTGCTCCGACACCCGCTCCGCTCC 1697
RESULT 4
AAV48324 AAV48324 standard; DNA; 60 BP.
XX
AC AAV48324;
XX
DT 20-NOV-1998 (first entry)
XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit R9.
XX
KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW tumour; tumour-associated antigen.
XX
OS Homo sapiens.
XX
PN WO9837095-A2.
XX
PD 27-AUG-1998.
XX
PF 24-FEB-1998; 98WO-US003693.
XX
PR 24-FEB-1997; 97US-0038253P.
XX
PA (THER-) THERION BIOLOGICS CORP.
(USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (DAND) DANA FARBER CANCER INST INC.
XX
PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX
DR WPI; 1998-467492/40.
XX
PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS Disclosure; Page 11; 42pp; English.
XX
XX The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ Sequence 60 BP; 12 A; 26 C; 14 G; 8 T; 0 U; 0 Other; .

Query Match 61.0%; Score 36.6; DB 2; Length 60;
Best Local Similarity 76.3%; Pred. No. 0.017;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

OY 1 GGGTCAACAGCTCTCCCGCTCATGGGGTTACTTGTCTCCAGATACTGCCAGCTCC 59
DB 1 GGTTCAACGGCAGCTCCAGCACAGGAGTCACTGTGCACCCGACACCCGCTCCAGCTCC 59

RESULT 5
AAV48323
ID AAV48323 standard; DNA; 60 BP.
XX
AC AAV48323;
XX
DT 20-NOV-1998 (first entry)
XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit R8.
XX
KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KM tumour; tumour-associated antigen.
XX
OS Homo sapiens.
XX
PN WO9837095-A2.
XX
PD 27-AUG-1998.
XX
PF 24-FEB-1998; 98WO-US003693.
XX
PR 24-FEB-1997; 97US-0038253P.
XX
PA (THER-) THERION BIOLOGICS CORP.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (DAND) DANA FARBER CANCER INST INC.
XX
PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX
DR WPI; 1998-467492/40.
XX
PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS Disclosure; Page 11; 42pp; English.
XX
XX The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an

CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ Sequence 60 BP; 10 A; 24 C; 17 G; 9 T; 0 U; 0 Other;

Query Match 60.0%; Score 36; DB 2; Length 60;
Best Local Similarity 75.0%; Pred. No. 0.026;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY 1 GGGTCAACAGCTCTCCCGCTCATGGGGTTACTTGTCTCCAGATACTGCCAGCTCCA 60
DB 1 GGGTCGAGTGGCCCTCCGGCGCATGGTGTGACCTCAGCTCTGACACAAAGGCCAGCCCA 60

RESULT 6
ADK68631
ID ADK68631 standard; cDNA; 120 BP.
XX
AC ADK68631;
XX
DT 06-MAY-2004 (first entry)
XX
DE HSP65-MUC1 antigen CTL epitope related cDNA #2.
XX
KW Human breast cancer; Mycobacterium bovis; heat shock protein 65; BCG;
KW HSP65; cytotoxic T lymphocyte; CTL; MUC1; HSP65-MUC1 antigen CTL epitope;
KW ss.
XX
OS Unidentified.
XX
PN CN1368384-A.
XX
PD 11-SEP-2002.
XX
PF 08-FEB-2001; 2001CN-00102614.
XX
PR 08-FEB-2001; 2001CN-00102614.
XX
PA (DIWE-) DIWEIHUAYU BIO TECHNOLOGY CO LTD BEIJING.
XX
PI Yu Y, Li H;
XX
DR WPI; 2003-854662/80.
XX
PT Genetically engineered vaccine of MUC-1 antigen for human breast cancer.
XX
PS Claim 4; SEQ ID NO 3; 14pp; Chinese.
XX
XX The invention relates to a method of preparation of a genetically
CC engineered vaccine for preventing and treating human breast cancer. The
CC method comprises fusing the coding gene of the Mycobacterium bovis heat
CC shock protein 65 (BCG HSP65) with the cytotoxic T lymphocyte (CTL)
CC epitope gene of MUC1 (antigen cell expressed by human breast cancer
CC cells) and the HSP65-MUC1 antigen CTL epitope is expressed in Escherichia
CC coli cells. This sequence represents DNA used in the method of the
XX invention.
XX
SQ Sequence 120 BP; 11 A; 43 C; 37 G; 27 T; 0 U; 2 Other;

Query Match 56.0%; Score 33.6; DB 10; Length 120;
Best Local Similarity 72.4%; Pred. No. 0.18;
Matches 42; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

OY 1 GGGTCAACAGCTCTCCCGCTCATGGGGTTACTTGTCTCCAGATACTGCCAGCTCC 58
DB 1 GGTTCTACCGGTCGCCNNCTCAGGGGTGTACTCTGTGCTCCGGCAACCGTCCGCTC 118

RESULT 7
AAV48318
ID AAV48318 standard; DNA; 60 BP.
XX
AC AAV48318;

```
XX
DT 20-NOV-1998 (first entry)
XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit R3.
XX
KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW tumour; tumour-associated antigen.
XX
OS Homo sapiens.
XX
PN WO9837095-A2.
XX
PD 27-AUG-1998.
XX
PF 24-FEB-1998; 98WO-US003693.
XX
PR 24-FEB-1997; 97US-0038253P.
XX
PA (THER-) THERION BIOLOGICS CORP.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA (DAND ) DANA FARBER CANCER INST INC.
XX
PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX
DR WPI; 1998-467492/40.
XX
PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS Disclosure; Page 11; 42pp; English.
XX
CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ Sequence 60 BP; 7 A; 29 C; 19 G; 5 T; 0 U; 0 Other;

Query Match 53.0%; Score 31.8; DB 2; Length 60;
Best Local Similarity 71.2%; Pred. No. 0.62;
Matches 42; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 GGGTCACAGCTCTCCCGCTCATGGGGTTACTTGTCTCCAGATTAATCGCCCACTCC 59
    |||||
Db 1 GGATCCACCGCGCGCGCTGCGACAGAGTAGACGTGCGCGCCGACACAGCGCCCGCTCC 59

RESULT 8
AAD00385
ID AAD00385 standard; DNA; 525 BP.
XX
AC AAD00385;
XX
DT 29-AUG-2000 (first entry)
XX
DE Human Mucin 1 (MUC-1) protein fragment encoding DNA #2.
XX
KW Human; Mucin 1; tumour; pMRS30 expression vector; anti-tumour;
KW therapy; immune response; cytostatic; vaccine; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..525
FT /tag= a
FT /product= "MUC-1 protein fragment"
XX
PN WO200025827-A2.
XX
```

```
PD 11-MAY-2000.
XX
XX 18-OCT-1999; 99WO-EP007874.
PF 30-OCT-1998; 98IT-MI002330.
XX
PR (MENA ) MENARINI RICERCHE SPA.
XX
PA Parente D, Di Massimo AM, De Santis R;
XX
PI WPI; 2000-365410/31.
XX
DR P-PSDB; AAY71021.
DR
XX
PT Composition containing one or more DNA molecules encoding fragments of a
PT Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
PT tumor therapy.
XX
PS Claim 16; Fig 2; 56pp; English.
XX
CC The present sequence is a DNA encoding a fragment of human Mucin 1 (MUC-
CC 1) antigenic protein which is overexpressed in tumour cells. The sequence
CC was obtained from BT20 tumour cells by reverse transcriptase- PCR and
CC corresponds to nucleotides 205-720 of the EMBL sequence J05581 with a
CC start codon and two stop codons. The present sequence is cloned into a
CC pMRS30 expression vector and used in pharmaceutical composition e.g.
CC vaccine for inducing an antigen-specific anti-tumour immune response.
CC Composition containing this DNA molecule is useful in anti-tumour therapy
CC of patients affected with tumours characterised by high MUC-1 expression
XX
SQ Sequence 525 BP; 111 A; 199 C; 128 G; 87 T; 0 U; 0 Other;

Query Match 52.7%; Score 31.6; DB 3; Length 525;
Best Local Similarity 74.1%; Pred. No. 1;
Matches 40; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 7 ACAGCTCTCCCGCTCATGGGGTTACTTGTCTCCAGATTAATCGCCCACTCCA 60
    |||||
Db 262 ACTGCTCCACGACACACAGGTGTACTCTCGGCTCCGATACCAAGCGGCCCA 315

RESULT 9
AAD00391
ID AAD00391 standard; DNA; 891 BP.
XX
AC AAD00391;
XX
DT 15-SEP-2003 (revised)
DT 29-AUG-2000 (first entry)
XX
DE Ubiquitin-E. coli Laci-human Mucin 1 fusion protein encoding DNA #2.
XX
KW Ubiquitin; Laci; beta-galactosidase; fusion protein; human; Mucin 1;
KW MUC-1; tumour; pMRS30 expression vector; anti-tumour; therapy;
KW immune response; cytostatic; vaccine; ds.
XX
OS Homo sapiens.
OS Escherichia coli.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 1..891
FT /tag= a
FT /product= "Ubiquitin-laci-MUC-1 fusion protein"
FT misc_feature 1..369
FT /tag= b
FT /label= UBILACI DNA
FT /note= "Includes ubiquitin-E. coli Laci fusion DNA"
FT 370..891
FT /tag= C
FT /note= "Human MUC-1 partial DNA that corresponds to
FT nucleotides 205-720 of the EMBL sequence J05581 with two
FT stop codons"
XX
```

```
PN WO200025827-A2.
XX
XX 11-MAY-2000.
PD
XX 18-OCT-1999; 99WO-EP007874.
PF
XX 30-OCT-1998; 98IT-MI002330.
PR
XX (MENA ) MENARINI RICERCHE SPA.
PA
XX
XX Parente D, Di Massimo AM, De Santis R;
PI
XX WPI; 2000-365410/31.
DR P-PSDB; AAY71027.
XX
XX Composition containing one or more DNA molecules encoding fragments of a
PT Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
PT tumor therapy.
XX
XX Claim 18; Fig 8; 56pp; English.
PS
XX The present sequence is a DNA encoding a fusion protein consisting of
CC human Mucin 1 (MUC-1) fragment fused to Ubilacti sequence at the N-
CC terminus. The Ubilacti sequence consists of ubiquitin from MCF7 cell line
CC and a portion of E. coli beta-galactosidase (LacI). MUC-1 is an antigenic
CC protein overexpressed in tumor cells. The present sequence is cloned
CC into a pMRS30 expression vector and used in pharmaceutical composition
CC e.g. vaccine for inducing an antigen-specific anti-tumour immune
CC response. Composition containing this DNA molecule is useful in anti-
CC tumour therapy of patients affected with tumours characterised by high
CC MUC-1 expression. (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ Sequence 891 BP; 203 A; 307 C; 225 G; 156 T; 0 U; 0 Other;

Query Match 52.7%; Score 31.6; DB 3; Length 891;
Best Local Similarity 74.1%; Pred. No. 1.1;
Matches 40; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 7 ACAGCTCTCCCGCTCATGGGGTTACTTCTGCTCCAGATACTCGCCAGCTCCA 60
DB 628 ACTGCTCCACGACACGCGTGTACTCTCGCTCCGATACGAGCGGCCCA 681

RESULT 10
AAD00388
ID AAD00388 standard; DNA; 1371 BP.
XX
XX AAD00388;
AC
XX 29-AUG-2000 (first entry)
DT
XX
XX Human Mucin 1 (MUC-1) protein fragment encoding DNA #5.
DE
XX Human; Mucin 1; tumour; pMRS30 expression vector; anti-tumour;
KM therapy; immune response; cytostatic; vaccine; ds.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH CDS 1..1371
FT /*tag= a
FT /product= "MUC-1 protein fragment"
FT
XX WO200025827-A2.
XX
XX 11-MAY-2000.
XX
XX 18-OCT-1999; 99WO-EP007874.
XX
XX 30-OCT-1998; 98IT-MI002330.
XX
XX (MENA ) MENARINI RICERCHE SPA.
XX
```

```
PI Parente D, Di Massimo AM, De Santis R;
XX
XX WPI; 2000-365410/31.
DR P-PSDB; AAY71024.
XX
XX Composition containing one or more DNA molecules encoding fragments of a
PT Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
PT tumor therapy.
XX
XX Claim 16; Fig 5; 56pp; English.
PS
XX The present sequence is a DNA encoding a fragment of human Mucin 1 (MUC-
CC 1) antigenic protein which is overexpressed in tumour cells. The sequence
CC was obtained by PCR from plasmids pMRS166, pMRS167, pMRS168 and pMRS169
CC which contain MUC-1 DNA from BT20 tumour cells. It corresponds to
CC nucleotides 136-1497 of the EMBL sequence J05581 with a start codon and
CC two stop codons. The present sequence is cloned into a pMRS30 expression
CC vector and used in pharmaceutical composition e.g. vaccine for inducing
CC an antigen-specific anti-tumour immune response. Composition containing
CC this DNA molecule is useful in anti-tumour therapy of patients affected
CC with tumours characterised by high MUC-1 expression
XX
SQ Sequence 1371 BP; 300 A; 450 C; 321 G; 300 T; 0 U; 0 Other;

Query Match 52.7%; Score 31.6; DB 3; Length 1371;
Best Local Similarity 74.1%; Pred. No. 1.2;
Matches 40; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 7 ACAGCTCTCCCGCTCATGGGGTTACTTCTGCTCCAGATACTCGCCAGCTCCA 60
DB 331 ACCGCTCCACGACACGCGTGTACTCTCGCTCCGATACGAGCGGCCCA 384

RESULT 11
AAD00394
ID AAD00394 standard; DNA; 1737 BP.
XX
XX AAD00394;
AC
XX 15-SEP-2003 (revised)
DT 29-AUG-2000 (first entry)
DT
XX
XX Ubiquitin-E. coli LacI-human Mucin 1 fusion protein encoding DNA #5.
DE
XX Ubiquitin; LacI; beta-galactosidase; fusion protein; human; Mucin 1;
KM MUC-1; tumour; pMRS30 expression vector; anti-tumour; therapy;
KM immune response; cytostatic; vaccine; ds.
XX
XX Homo sapiens.
OS
XX Escherichia coli.
OS
XX Chimeric.
XX
XX Key Location/Qualifiers
FH CDS 1..1737
FT /*tag= a
FT /product= "Ubiquitin-LacI-MUC-1 fusion protein"
FT
XX misc_feature 1..369
FT /*tag= b
FT /label= UBILACTI_DNA
FT /note= "Includes ubiquitin-E. coli LacI fusion DNA"
FT 370..1737
FT /*tag= c
FT /note= "Human MUC-1 partial DNA that corresponds to
FT nucleotides 136-1497 of the EMBL sequence J05581 with two
FT stop codons"
XX
XX WO200025827-A2.
XX
XX 11-MAY-2000.
XX
XX 18-OCT-1999; 99WO-EP007874.
XX
XX 30-OCT-1998; 98IT-MI002330.
XX
```

```
XX (MENA ) MENARINI RICERCHE SPA.
PA
XX Parente D, Di Massimo AM, De Santis R;
PI
XX WPI; 2000-365410/31.
DR P-PSDB; AAY71030.
XX
XX Composition containing one or more DNA molecules encoding fragments of a
PT Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
PT tumor therapy.
XX
XX Claim 18; Fig 11; 56pp; English.
XX
XX The present sequence is a DNA encoding a fusion protein consisting of
CC human Mucin 1 (MUC-1) fragment fused to UBIlaci sequence at the N-
CC terminus. The UBIlaci sequence consists of ubiquitin from MCF7 cell line
CC and a portion of E. coli beta-galactosidase (Laci). MUC-1 is an antigenic
CC protein overexpressed in tumour cells. The present sequence is cloned
CC into a pMR530 expression vector and used in pharmaceutical composition
CC e.g. vaccine for inducing an antigen-specific anti-tumour immune
CC response. Composition containing this DNA molecule is useful in anti-
CC tumour therapy of patients affected with tumours characterised by high
CC MUC-1 expression. (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ Sequence 1737 BP; 392 A; 558 C; 418 G; 369 T; 0 U; 0 Other;

Query Match          52.7%; Score 31.6; DB 3; Length 1737;
Best Local Similarity 74.1%; Pred. No. 1.3;
Matches 40; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 7 ACAGCTCCTCCCGCTCATGGGGTTACTTCTGCTCCAGATACTCGCCAGCTCCA 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 697 ACCGCTCCACCAGACACGCGTGTACTCTCGCTCCGATACACAGCGCGGCCCA 750

RESULT 12
AAV48320
ID AAV48320 standard; DNA; 60 BP.
XX
AC AAV48320;
XX
DT 20-NOV-1998 (first entry)
XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit R5.
XX
DE ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW tumour; tumour-associated antigen.
XX
OS Homo sapiens.
XX
PN WO9837095-A2.
XX
PD 27-AUG-1998.
XX
PF 24-FEB-1998; 98WO-US003693.
XX
PR 24-FEB-1997; 97US-0038253P.
XX
PA (THER-) THERION BIOLOGICS CORP.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA (DAND ) DANA FARBER CANCER INST INC.
XX
PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX
DR WPI; 1998-467492/40.
XX
PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS Disclosure; Page 11; 42pp; English.
XX
CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
```

```
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ Sequence 60 BP; 9 A; 27 C; 15 G; 9 T; 0 U; 0 Other;

Query Match          50.3%; Score 30.2; DB 2; Length 60;
Best Local Similarity 69.5%; Pred. No. 2.1;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGGTCACACACTCCTCCCGCTCATGGGTTACTTCTGCTCCAGATACTCGCCAGCTCC 59
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 GGTTCGACGCGCCCGCTGCTCAGCGGTATACATCCGCCCCGATACCAAGACCGCCCC 59

RESULT 13
AAS01385
ID AAS01385 standard; DNA; 68 BP.
XX
AC AAS01385;
XX
DT 04-JUL-2001 (first entry)
XX
DE SBMV coat protein-MUC(16) insertion construct #3.
XX
KW Human; polymorphic epithelial cell mucin; PEM; chimeric virus particle;
KW CVP; plant virus coat protein; comovirus; CPMV; cowpea mosaic virus;
KW SBMV; Southern bean mosaic virus; LTSV; red clover necrotic mosaic virus;
KW RCNMV; MUC1; tumour; cancer vaccine; mutant; ds.
XX
OS Southern bean mosaic virus.
OS Homo sapiens.
OS Synthetic.
OS Chimeric.
XX
XX Key Location/Qualifiers
FH misc_recomb 12..59
FT /*tag= a
FT /note= "Human MUC(16) sequence (AAS01380)"
XX
PN WO200118199-A1.
XX
PD 15-MAR-2001.
XX
PF 11-SEP-2000; 2000WO-GB003500.
XX
PR 09-SEP-1999; 99GB-00021337.
XX
PA (DOWC ) DOW CHEM CO.
XX
PI Bendig MM, Jones TD, Longstaff M, Hellendoorn K;
XX
DR WPI; 2001-244570/25.
XX
XX New chimeric plant virus particles with an immunogenically active peptide
PT of a tumor-associated mucin, useful as a vaccine or for the manufacture
PT of a vaccine for treating and/or preventing tumors and/or cancer.
XX
XX Example 12; Page; 63pp; English.
XX
XX The present sequence for SBMV (Southern bean mosaic virus) coat protein-
CC MUC(16) insertion construct #3 is 1 of 5 constructs containing a MUC1(16)
CC DNA sequence at different positions within a SBMV coat protein insertion
CC sequence (AAS01382). The construct is used to create a novel chimeric
CC virus particle (CVP). MUC1(16) peptide is 1 of 7 polymorphic epithelial
CC cell mucin (PEM) peptides (AAU00483-AAU00489) which can be used to create
CC CVPs. The mucin peptide epitopes are preferably inserted into the coat
CC protein of a plant virus such as the comovirus CPMV (cowpea mosaic
CC virus). The ability of the chimeric virus particle CPMV-MUC1(16) to
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 17:40:52 ; Search time 722.8 Seconds
(without alignments)
3925.542 Million cell updates/sec

Title: US-10-057-136-6

Perfect score: 60

Sequence: 1 gggctcacagctcctcccgccg.....cagatactcgcccgctcca 60

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	2297	6 BD272907	BD272907 A recombi
2	33.6	56.0	176263	2 AC102389	AC102389 Mus muscu
3	33.6	56.0	223290	2 AC140409	AC140409 Mus muscu
4	31.6	52.7	525	6 BD225141	BD225141 Medicinal
5	31.6	52.7	891	6 BD225147	BD225147 Medicinal
6	31.6	52.7	1371	6 BD225144	BD225144 Medicinal
7	31.6	52.7	1737	6 BD225150	BD225150 Medicinal
8	30.8	51.3	149500	2 AC138636	AC138636 Leishmani
9	29.6	49.3	93	6 AX703426	AX703426 Sequence
10	29.6	49.3	157	6 AX703428	AX703428 Sequence
11	29.6	49.3	82400	10 AC090495	AC090495 Genomic B
12	29.6	49.3	164629	10 AC121094	AC121094 Mus muscu
13	29.6	49.3	174264	10 AL591373	AL591373 Mouse DNA
14	29.6	49.3	207182	10 AC144938	AC144938 Mus muscu
15	29.6	49.3	219663	2 AC124587	AC124587 Mus muscu
16	29.6	49.3	223724	10 AC114003	AC114003 Mus muscu
17	29	48.3	1638	10 BC026670	BC026670 Mus muscu
18	29	48.3	2446	4 OC037769	U37769 Oryctolagus
19	29	48.3	2524	10 BC058977	BC058977 Mus muscu

C	20	28.8	48.0	249767	10 AC114539	AC114539 Mus muscu
	21	28.6	47.7	519	6 AR316604	AR316604 Sequence
	22	28.6	47.7	534	6 AR316608	AR316608 Sequence
	23	28.6	47.7	40666	2 AC019903	AC019903 Drosophil
	24	28.6	47.7	86391	3 AC004716	AC004716 Drosophil
	25	28.6	47.7	171000	3 AC092244	AC092244 Drosophil
	26	28.6	47.7	192707	2 AC116250	AC116250 Rattus no
	27	28.6	47.7	195798	2 AC118890	AC118890 Rattus no
	28	28.6	47.7	337636	3 AE003584	AE003584 Drosophil
	29	28.4	47.3	110000	8 CR382131_07	Continuation (8 of
	30	28.2	47.0	116409	10 AC131118	AC131118 Mus muscu
	31	28.2	47.0	178544	10 AC124557	AC124557 Mus muscu
	32	28.2	47.0	213037	2 AC093986	AC093986 Rattus no
	33	28.2	47.0	233411	2 AC109951	AC109951 Rattus no
	34	28.2	47.0	234248	2 AC119385	AC119385 Rattus no
	35	28.2	47.0	259308	2 AC111121	AC111121 Mus muscu
	36	28	46.7	110000	2 AC091341_5	Continuation (6 of
	37	28	46.7	124244	10 AL929026	AL929026 Mouse DNA
	38	28	46.7	144093	10 AC109193	AC109193 Mus muscu
	39	28	46.7	157092	10 AL805911	AL805911 Mouse DNA
	40	28	46.7	194209	10 AL671894	AL671894 Mouse DNA
	41	28	46.7	199993	10 AC102783	AC102783 Mus muscu
	42	28	46.7	200574	10 BX005304	BX005304 Mouse DNA
	43	28	46.7	202311	2 AC116769	AC116769 Mus muscu
	44	28	46.7	205054	10 AC098712	AC098712 Mus muscu
	45	28	46.7	225238	2 AC097825	AC097825 Rattus no

ALIGNMENTS

RESULT 1	BD272907	2297 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD272907	A recombinant vector expressing multiple constitutlatory molecules and uses thereof.			
DEFINITION	BD272907	A recombinant vector expressing multiple constitutlatory molecules and uses thereof.			
ACCESSION	BD272907	GI:33082675			
VERSION	BD272907.1	JP 2002531133-A/1.			
KEYWORDS	JP 2002531133-A/1.	synthetic construct			
SOURCE	JP 2002531133-A/1.	artificial sequences.			
ORGANISM	JP 2002531133-A/1.	Schlom,J., Hodge,J. and Panicali,D.			
REFERENCE	1 (bases 1 to 2297)	A recombinant vector expressing multiple constitutlatory molecules and uses thereof			
AUTHORS	Schlom,J., Hodge,J. and Panicali,D.	Patent: JP 2002531133-A 1 24-SEP-2002;			
TITLE	THE UNITED STATES OF AMERICA,THERION BIOLOGICS CORP	OS Artificial Sequence			
JOURNAL	JP 2002531133-A/1	PD 24-SEP-2002			
COMMENT	PF 12-NOV-1999 JP 2000586927	PR 09-DEC-1998 US 60/111582			
	PI JEFFREY SCHLOM,JAMES HODGE,DENNIS PANICALI	PC C12N15/02,A61K35/12,A61K35/74,A61K35/76,A61K38/00,A61K39/00,A61K39/12,A61K39/12,A61K39/21,A61K39/235,A61K39/245,A61K39/275,A61K39/29,A61K48/00,A61P1/04,A61P31/04,A61P31/10,A61P31/12,A61P35/00,A61P37/02, PC A61P37/06,			
	PC C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N7/00,C12Q1/02,G01N33/53,C12N15/00,	PC C12N5/00,A61K37/02			
	CC Description of Artificial Sequence: VECTOR SEQUENCE FH	Key			
	FT source	1. .2297			
FEATURES	Location/Qualifiers	/organism='Artificial Sequence'.			
source	1. .2297	/organism="synthetic construct"			
		/mol_type="genomic DNA"			
		/db_xref="taxon:32630"			
ORIGIN					

Query Match 100.0%; Score 60; DB 6; Length 2297;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCAACAGCTCTCCCGCTCATGGGGTACTTCTGCTCCAGATACTCGCCAGCTCCA 60
|||||
526 GGCTCAACAGCTCTCCCGCTCATGGGGTACTTCTGCTCCAGATACTCGCCAGCTCCA 585

RESULT 2
AC102389/c
LOCUS
DEFINITION Mus musculus clone RP24-374N16, WORKING DRAFT SEQUENCE, 12
AC102389
unordered pieces.
AC102389
AC102389.3 GI:28893769
HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 176263)
Birtren,B., Nusbaum,C. and Lander,E.
Mus musculus, clone RP24-374N16
Unpublished
2 (bases 1 to 176263)
Birtren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczyk,J., Levine,R., Liu,G., Maclean,C., Macdonald,P., Major,J., Marguis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Menus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 176263)
Birtren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Menus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupack,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL
REFERENCE
AUTHORS

TITLE
JOURNAL
COMMENT

Direct Submission
Submitted (09-MAR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 9, 2003 this sequence version replaced gi:22381003.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L18673
Center clone name: 374 N 16
Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 173405 bases at least Q40
Consensus quality: 174612 bases at least Q30
Consensus quality: 174923 bases at least Q20
Insert size: 170000; agarose-fp
Insert coverage: 175163; sum-of-contigs
Quality coverage: 9.5 in Q20 bases; agarose-fp
Quality coverage: 9.2 in Q20 bases; sum-of-contigs
NOTE: This is a 'working draft' sequence. It currently consists of 12 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1859: contig of 1859 bp in length
* 1860 1959: gap of 100 bp
* 1960 3317: contig of 1358 bp in length
* 3318 3417: gap of 100 bp
* 3418 4588: contig of 1171 bp in length
* 4589 4688: gap of 100 bp
* 4689 6418: contig of 1730 bp in length
* 6419 6518: gap of 100 bp
* 6519 10148: contig of 3630 bp in length
* 10149 10248: gap of 100 bp
* 10249 15376: contig of 5128 bp in length
* 15377 15476: gap of 100 bp
* 15477 26015: contig of 10539 bp in length
* 26016 26115: gap of 100 bp
* 26116 40612: contig of 14497 bp in length
* 40613 40712: gap of 100 bp
* 40713 58063: contig of 17351 bp in length
* 58064 58163: gap of 100 bp
* 58164 82401: contig of 24238 bp in length
* 82402 82501: gap of 100 bp
* 82502 140587: contig of 58086 bp in length
* 140588 140687: gap of 100 bp
* 140688 176263: contig of 35576 bp in length.
Location/Qualifiers
1. 176263
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP24-374N16"
/clone_lib="RPCI-24 Male Mouse BAC"
1. 1859
/note="assembly_fragment
clone_end:SP6
vector_side:left"
1960. 3317
/note="assembly_fragment"
3418. 4588
/note="assembly_fragment"

FEATURES
source
misc_feature
misc_feature
misc_feature

misc_feature 4689. .6418 /note="assembly_fragment"
misc_feature 6519. .10148 /note="assembly_fragment"
misc_feature 10249. .15376 /note="assembly_fragment"
misc_feature 15477. .26015 /note="assembly_fragment"
misc_feature 26116. .40612 /note="assembly_fragment"
misc_feature 40713. .58063 /note="assembly_fragment"
misc_feature 58164. .82401 /note="assembly_fragment"
misc_feature 82502. .140587 /note="assembly_fragment"
misc_feature 140688. .176263 /note="assembly_fragment"
misc_feature /note="assembly_fragment"
clone_end:T7
vector_side:right"

ORIGIN
Query Match 56.0%; Score 33.6; DB 2; Length 176263;
Best Local Similarity 75.0%; Pred. No. 1.9;
Matches 42; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Oy 5 CAACAGCTCCTCCCGCTCATGGGTTACTTCTGCTCCAGATCTGCCAGCTCCA 60
Db 36031 CAGTGCTCCCCCAGCTCCAGGTGTCTCCCGCAGCTCCAGATGCTCCCGCAGCTCCA 35976

RESULT 3
AC140409 223290 bp DNA linear HTG 23-FEB-2003
LOCUS Mus musculus chromosome UNK clone RP23-138K22, WORKING DRAFT
DEFINITION
SEQUENCE, 16 unordered pieces.

AC140409 AC140409.1 GI:28475673
VERSION HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 223290)
JOURNAL McPherson,J.D. and Waterston,R.H.
REFERENCE The sequence of Mus musculus clone
AUTHORS Unpublished
TITLE 2 (bases 1 to 223290)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (23-FEB-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA

COMMENT

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
Contact: submissions@watson.wustl.edu
----- Project Information -----
Center project name: M_BA0138K22

----- Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing vector: plasmid; 100%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 218110 bases at least Q40
Consensus quality: 218643 bases at least Q30
Consensus quality: 219019 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces

* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1248: contig of 1248 bp in length
1249 1348: gap of unknown length
1349 4137: contig of 2789 bp in length
4138 4237: gap of unknown length
4238 8834: contig of 4597 bp in length
8835 8935: gap of unknown length
8935 14846: contig of 5912 bp in length
14847 14946: gap of unknown length
14947 23063: contig of 8117 bp in length
23064 23163: gap of unknown length
23164 32288: contig of 9125 bp in length
32289 32388: gap of unknown length
32389 44729: contig of 12341 bp in length
44730 44829: gap of unknown length
44830 61526: contig of 16697 bp in length
61527 61626: gap of unknown length
61627 72386: contig of 10760 bp in length
72387 72486: gap of unknown length
72487 85071: contig of 12585 bp in length
85072 85171: gap of unknown length
85172 98342: contig of 13171 bp in length
98343 98442: gap of unknown length
98443 114309: contig of 15867 bp in length
114310 114409: gap of unknown length
114410 141127: contig of 26718 bp in length
141128 141227: gap of unknown length
141228 165507: contig of 24280 bp in length
165508 165607: gap of unknown length
165608 194683: contig of 29076 bp in length
194684 223290: contig of unknown length
194784 223290: contig of 28507 bp in length.

FEATURES
source Location/Qualifiers

1. .223290
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="UNK"
/clone="RP23-138K22"
1. .1248
/note="assembly_name:Contig15"
1349. .4137
/note="assembly_name:Contig18"
4238. .8834
/note="assembly_name:Contig19"
8935. .14846
/note="assembly_name:Contig20"
14947. .23063
/note="assembly_name:Contig21"
23164. .32288
/note="assembly_name:Contig22"
32389. .44729
/note="assembly_name:Contig23"
44830. .61526
/note="assembly_name:Contig24"
61627. .72386
/note="assembly_name:Contig25"
72487. .85071
/note="assembly_name:Contig26"
85172. .98342
/note="assembly_name:Contig27"
98443. .114309
/note="assembly_name:Contig28"
114410. .141127
/note="assembly_name:Contig29"
141228. .165507
/note="assembly_name:Contig30"
165608. .194683

misc_feature 194784..223290 /note="assembly_name:Contig31"
/note="assembly_name:Contig32"

ORIGIN

Query Match 56.0%; Score 33.6; DB 2; Length 223290;
Best Local Similarity 75.0%; Pred. No. 1.9;
Matches 42; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 5 CAACAGCTCCTCCCGCTCATGGGGTACTTGTCTCCAGATACTCGCCAGCTCCA 60
Db 111887 CAGTGCTCTCCCGCAGCTCCAGGTGTTCCTCCAGCTCCAGATGCTCCCGCAGCTCCA 111942

RESULT 4
BD225141 525 bp DNA linear PAT 17-JUL-2003
DEFINITION Medicinal composition having antitumor effect and containing DNA
ACCESSION BD225141
VERSION BD225141.1 GI:33034911
KEYWORDS JP 2002528519-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 525)
Pallente,D., Massimo,A.M.D. and Desantis,R.
Medicinal composition having antitumor effect and containing DNA
encoding antigenic protein
Patent: JP 2002528519-A 2 03-SEP-2002;
MENARINI RICERCHE SPA

COMMENT OS Homo sapiens (human)
PN JP 2002528519-A/2
PD 03-SEP-2002
PF 18-OCT-1999 JP 2000579265
PR 30-OCT-1998 IT MI98A002330
PI DINO PALLENTE,ANNA MARIA D MASSIMO,RITA DESANTIS PC
A61K38/00,A61K35/76,A61K39/00,A61K48/00,A61P35/00,C12N15/09, PC
A61K37/02,
PC C12N15/00
CC Medicinal composition having antitumor effect and containing
CC DNA encoding
CC antigenic protein
CC key
FH Key Location/Qualifiers
FT source 1..525
FEATURES source location/Qualifiers
1..525
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 52.7%; Score 31.6; DB 6; Length 525;
Best Local Similarity 74.1%; Pred. No. 5.4;
Matches 40; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 7 ACAGTCTCTCCCGCTCATGGGGTACTTGTCTCCAGATACTCGCCAGCTCCA 60
Db 262 ACTGCTCCACGACACGCGTGTACTCTGGCTCCGATACCGAGCGCGCCCA 315

RESULT 5
BD225147 891 bp DNA linear PAT 17-JUL-2003
DEFINITION Medicinal composition having antitumor effect and containing DNA
ACCESSION BD225147
VERSION BD225147.1 GI:33034917
KEYWORDS JP 2002528519-A/8.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE AUTHORS
TITLE
JOURNAL
COMMENT OS Homo sapiens (human)
PN JP 2002528519-A/8
PD 03-SEP-2002
PF 18-OCT-1999 JP 2000579265
PR 30-OCT-1998 IT MI98A002330
PI DINO PALLENTE,ANNA MARIA D MASSIMO,RITA DESANTIS PC
A61K38/00,A61K35/76,A61K39/00,A61K48/00,A61P35/00,C12N15/09, PC
A61K37/02,
PC C12N15/00
CC Medicinal composition having antitumor effect and containing
CC DNA encoding
CC antigenic protein
CC key
FH Key Location/Qualifiers
FT source 1..891
FEATURES source location/Qualifiers
1..891
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 52.7%; Score 31.6; DB 6; Length 891;
Best Local Similarity 74.1%; Pred. No. 5.6;
Matches 40; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 7 ACAGTCTCTCCCGCTCATGGGGTACTTGTCTCCAGATACTCGCCAGCTCCA 60
Db 628 ACTGCTCCACGACACGCGTGTACTCTCGGCTCCGATACCGAGCGCGCCCA 681

RESULT 6
BD225144 1371 bp DNA linear PAT 17-JUL-2003
DEFINITION Medicinal composition having antitumor effect and containing DNA
ACCESSION BD225144
VERSION BD225144.1 GI:33034914
KEYWORDS JP 2002528519-A/5.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1371)
Pallente,D., Massimo,A.M.D. and Desantis,R.
Medicinal composition having antitumor effect and containing DNA
encoding antigenic protein
Patent: JP 2002528519-A 5 03-SEP-2002;
MENARINI RICERCHE SPA

COMMENT OS Homo sapiens (human)
PN JP 2002528519-A/5
PD 03-SEP-2002
PF 18-OCT-1999 JP 2000579265
PR 30-OCT-1998 IT MI98A002330
PI DINO PALLENTE,ANNA MARIA D MASSIMO,RITA DESANTIS PC
A61K38/00,A61K35/76,A61K39/00,A61K48/00,A61P35/00,C12N15/09, PC
A61K37/02,
PC C12N15/00
CC Medicinal composition having antitumor effect and containing
CC DNA encoding
CC antigenic protein
CC key
FH Key Location/Qualifiers
FT source 1..1371
FEATURES source location/Qualifiers
1..1371
/organism="Homo sapiens (human)"

```

SOURCE                                1..1371
                                      /organism="Homo sapiens"
                                      /mol_type="genomic DNA"
                                      /db_xref="taxon:9606"
ORIGIN
Query Match                        52.7%; Score 31.6; DB 6; Length 1371;
Best Local Similarity 74.1%; Pred. No. 5.8;
Matches 40; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 7 ACAGCTCCTCCCGCTCATGGGGTAACTTGTGCTCCAGATACCTCGCCAGCTCCA 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 331 ACCGCTCCACCAGACACAGGTGTAACTCTCGGCTCCGGATACCAGCGCCGCCCA 384

RESULT 7
BD225150                        1737 bp   DNA       linear   PAT 17-JUL-2003
LOCUS
DEFINITION
Medicinal composition having anticancer effect and containing DNA
encoding antigenic protein.
BD225150
BD225150.1 GI:33034920
JP 2002528519-A/11.
Homo sapiens (human)
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1737)
Pallente, D., Massimo, A.M.D. and Desantis, R.
Medicinal composition having anticancer effect and containing DNA
encoding antigenic protein
Patent: JP 2002528519-A 11 03-SEP-2002;
MENARINI RICERCHES SPA
OS Homo sapiens (human)
PN JP 2002528519-A/11
PD 03-SEP-2002
PF 18-OCT-1999 JP 2000579265
PR 30-OCT-1998 IT MI98A002330
PI DINO PALLENTI, ANNA MARIA D MASSIMO, RITA DESANTIS PC
A61K38/00, A61K35/76, A61K39/00, A61K48/00, A61P35/00, C12N15/09, PC
A61K37/02,
PC C12N15/00
CC Medicinal composition having anticancer effect and containing
CC DNA encoding
CC antigenic protein
CC Key Location/Qualifiers
FH source 1..1737
FT Location/Qualifiers
1..1737
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
FEATURES
source
1..1737
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match                        52.7%; Score 31.6; DB 6; Length 1737;
Best Local Similarity 74.1%; Pred. No. 6;
Matches 40; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 7 ACAGCTCCTCCCGCTCATGGGGTAACTTGTGCTCCAGATACCTCGCCAGCTCCA 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 697 ACCGCTCCACCAGACACAGGTGTAACTCTCGGCTCCGGATACCAGCGCCGCCCA 750

RESULT 8
AC138636                        149500 bp   DNA       linear   HTG 25-FEB-2003
LOCUS
DEFINITION
Leishmania major chromosome 29 clone LB00573 strain Friedlin, ***
SEQUENCING IN PROGRESS ***, 6 ordered pieces.
AC138636
AC138636.2 GI:28557887
HTG: HTGS_PHASE2.
Leishmania major
SOURCE

```

```

ORGANISM      Leishmania major
REFERENCE      Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae;
AUTHORS        Leishmania.
TITLE          1 (bases 1 to 149500)
JOURNAL        Myler,P.J., Sisk,E., Vogt,C., Robertson,L., Munden,H., Worthey,L.,
AUTHORS        Nelson,S., Ivens,A., Seyler,A., Rinta,J. and Stuart,K.
REFERENCE      Submitted Submission
AUTHORS        Submitted (14-JAN-2003) Seattle Biomedical Research Institution, 4
JOURNAL        Nickerson Street, Seattle, WA 98109-1651, USA
TITLE          2 (bases 1 to 149500)
AUTHORS        Myler,P.J., Sisk,E., Vogt,C., Robertson,L., Munden,H., Worthey,L.,
REFERENCE      Nelson,S., Ivens,A., Seyler,A., Rinta,J. and Stuart,K.
JOURNAL        Direct Submission
COMMENT        Submitted (25-FEB-2003) Seattle Biomedical Research Institution, 4
AUTHORS        Nickerson Street, Seattle, WA 98109-1651, USA
TITLE          On Feb 25, 2003 this sequence version replaced gi:27733947.
JOURNAL        * NOTE: This is a 'working draft' sequence. It currently
AUTHORS        * consists of 6 contigs. Gaps between the contigs
REFERENCE      * are represented as runs of N. The order of the pieces
JOURNAL        * is believed to be correct as given, however the sizes
AUTHORS        * of the gaps between them are based on estimates that have
REFERENCE      * provided by the submittor.
JOURNAL        * This sequence will be replaced
AUTHORS        * by the finished sequence as soon as it is available and
REFERENCE      * the accession number will be preserved.
JOURNAL        1 17713: contig of 17713 bp in length
AUTHORS        * 17714 18016: gap of unknown length
REFERENCE      * 18017 21356: contig of 3340 bp in length
JOURNAL        * 21357 21658: gap of unknown length
AUTHORS        * 21659 26396: contig of 4738 bp in length
REFERENCE      * 26397 26698: gap of unknown length
JOURNAL        * 26699 72247: contig of 45549 bp in length
AUTHORS        * 72248 72549: gap of unknown length
REFERENCE      * 72550 123849: contig of 51300 bp in length
JOURNAL        * 123850 124151: gap of unknown length
AUTHORS        * 124152 149500: contig of 25349 bp in length.
FEATURES       Location/Qualifiers
SOURCE         1..149500
                /organism="Leishmania major"
                /mol_type="genomic DNA"
                /strain="Friedlin"
                /db_xref="taxon:5664"
                /chromosome="29"
                /clone="LB00573"

ORIGIN
Query Match          51.3%; Score 30.8; DB 2; Length 149500;
Best Local Similarity 70.7%; Pred. No. 16;
Matches 41; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY      2  GGTCAACAGCTCTCCCGCTCATGGGGTTACTTCTGCTCCAGATACCTCGCCAGCTCC 59
          |||||
Db       116463  GTGAACAGCTGTACTGCTCACGAGGCTGCTGCGCTGTCTCTCGACAACTGC 116520

RESULT 9
AX703426/c          93 bp      DNA      linear      PAT 03-APR-2003
LOCUS              Sequence 4 from Patent WO02086505.
DEFINITION         AX703426
ACCESSION          AX703426
VERSION            AX703426.1  GI:29538417
KEYWORDS
SOURCE             .
ORGANISM           synthetic construct
                  synthetic construct
                  artificial sequences.
REFERENCE          1
AUTHORS            Benson,R.S.
TITLE              Intracellular analysis
JOURNAL            Patent: WO 02086505-A 4 31-OCT-2002;
                  THE VICTORIA UNIVERSITY OF MANCHESTER (GB)
FEATURES           Location/Qualifiers
SOURCE             1..93
                  /organism="synthetic construct"

```


Landers, T., Lehoczký, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mithova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Nordu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission

JOURNAL Submitted (15-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 164629)

REFERENCE AUTHORS

Birken, B., Nusbbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearrellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fato, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Haasopian, D., Hacos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karataas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., McLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retra, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission

JOURNAL Submitted (24-APR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE 4 (bases 1 to 164629)

4 (bases 1 to 164629)

REFERENCE AUTHORS

Anderson, M., Anderson, S., Archchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslawski, L., Bouhgalter, B., Camarata, J., Chang, J., Choepele, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Doolley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunhchang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Teffaye, S., Theodore, J., Toplam, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission

JOURNAL Submitted (30-JUL-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Jul 30, 2004 this sequence version replaced gi:46518664.

On Jul 30, 2004 this sequence version replaced gi:46518664.

All repeats were identified using RepeatMasker:

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center

Center: Whitehead Institute/MIT Center for Genome Research
Center code: WIBR

Web site: <http://www-seg.wi.mit.edu>
 Contact: sequence_submissions@road.mit.edu

 Project Information
 Center project name: l21248
 Center clone name: 322_I_3

Some of the sequence contained within base pairs 119380 to the end of the clone was stolen from accession AC144938.

FEATURES

Source

```
1. 164629
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="18"
/map="18"
```

misc_feature

```
/note="clone_boundary
clone_end:SP6
site:MboI"
```

```
repeat_region complement(3. .179)
repeat_region /rpt_family="B4A"
repeat_region complement(1847. .2028)
```

```
/rpt_family="AT_ric
```

```
repeat_region /rpt_family="(TAA)n
5294. .5319
```

```
repeat_region 5797.5844  
/rpt_family="(TTTTTG)n"
```

```

repeat_region
7436.7559

```

```
repeat_region 7630..7660
               /rpt_family="(A)n"
```

```
repeat_region /rpt_family="(GA)n"
7871. .8023
```

```
repeat_region 8027..8070
/rpt_family=" (CGAG)
1015 1025
```

```

/rpt_family="At_rich"
11283. .11312
repeat_region

```

```
repeat_region 14721.14888
/rpt_family="CT-ric
sam1_start/14057
```

```

/note="single clone coverage
complement(14857. .14864)
/note="30 and single clone
unsure

```

```

unbre          complement(14884. .
recreat rection 14815 15039
              /note="<30 qual sim

```

```
repeat_region      16444. .16631
                    /rnt fam11v="(TG)n"
```

```

unbure      16785 /note="single clone coverage"
repeat region 16648

```

```

/1pc_family= (15)/n
17445. .17501
repeat_region
/rot_familyv="(GA)n"

```

```
repeat region
18040..18068
/xpt_family="(TC)n"
```

```
repeat_region complement(19489..19649)
```

```
repeat_region      /rpt_family="MLT1J2"
                    25239..25264
                    /rpt_family="(CA)n"
repeat_region      25346..25389
                    /rpt_family="(CA)n"
repeat_region      25390..25416
                    /rpt_family="(GA)n"
repeat_region      26450..26510
                    /rpt_family="(TG)n"
repeat_region      26926..26995
                    /rpt_family="(CAGAGA)n"
repeat_region      27650..27790
                    /rpt_family="L2"
repeat_region      27921..27954
                    /rpt_family="(CA)n"
repeat_region      28145..28311
                    /rpt_family="MIR"
repeat_region      29591..29696
                    /rpt_family="MIR"
repeat_region      29889
                    /rpt_family="(29731..
                    /rpt_family="MIR"
                    complement(30334..
                    /rpt_family="L2"
                    complement(32737..
                    /rpt_family="B4A"
                    33251..33481
                    /rpt_family="MIR"
                    33884..34054
                    /rpt_family="RSINE1"
                    complement(34315..
                    /rpt_family="B1F"
```

```
Query Match      49.3%; Score 29.6; DB 10; Length 164629;
Best Local Similarity 73.1%; Pred. No. 42;
Matches 38; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY      8 CAGCTCCTCCCGCTCATGGGGTTACTTGTGCTCCAGATACTGCCAGCTCC 59
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      153766 CAGCTCCTCCACCTCCTCCTGCTCCCTCCCTGCTCTCCACTCCCCAGCTCC 153715
```

```
RESULT 13
AL591373      174264 bp      DNA      linear      ROD 29-JUN-2002
LOCUS      Mouse DNA sequence from clone RP23-268M6 on chromosome 18, complete
DEFINITION
sequence.
ACCESSION      AL591373
VERSION      AL591373.8 GI:15983933
KEYWORDS      HTG.
SOURCE      Mus musculus (house mouse)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 174264)
AUTHORS      Phillimore,B.
TITLE      Direct Submission
JOURNAL      Submitted (29-JUN-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Oct 8, 2001 this sequence version replaced gi:15591499.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
```

abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-268M6 is from the RPCI-23 Mouse PAC Library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBACE3.6

----- Genome Center
Center: UK Medical Research Council
Center code: UK-MRC
Web site: <http://mrcseq.har.mrc.ac.uk>
Contact: mousegen@har.mrc.ac.uk

FEATURES

source

Location/Qualifiers
1..174264
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="18"
/clone="RP23-268M6"
/clone_lib="RPCI-23"

ORIGIN

```
Query Match      49.3%; Score 29.6; DB 10; Length 174264;
Best Local Similarity 73.1%; Pred. No. 42;
Matches 38; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
```

```
QY      8 CAGCTCCTCCCGCTCATGGGGTTACTTGTGCTCCAGATACTGCCAGCTCC 59
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      111108 CAGCTCCTCCAGCTCCTCCTCAGCTCCTCCTGCTCTCTCCAGCTCC 111159
```

```
RESULT 14
AC144938      207182 bp      DNA      linear      ROD 15-MAY-2004
LOCUS      Mus musculus BAC clone RP24-129H18 from chromosome 18, complete
DEFINITION
sequence.
ACCESSION      AC144938
VERSION      AC144938.3 GI:45237301
KEYWORDS      HTG.
SOURCE      Mus musculus (house mouse)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 207182)
AUTHORS      Sweenngen-Shahid,S., Shahid,S., Bielicki,L. and Haakenson,W.
TITLE      The sequence of Mus musculus BAC clone RP24-129H18
JOURNAL      Unpublished (2001)
REFERENCE      2 (bases 1 to 207182)
AUTHORS      Wilson,R.K.
TITLE      Direct Submission
JOURNAL      Submitted (27-MAY-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE      3 (bases 1 to 207182)
AUTHORS      Wilson,R.K.
TITLE      Direct Submission
JOURNAL      Submitted (25-JUN-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE      4 (bases 1 to 207182)
AUTHORS      Wilson,R.K.
TITLE      Direct Submission
JOURNAL      Submitted (06-MAR-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE      5 (bases 1 to 207182)
AUTHORS      Wilson,R.K.
TITLE      Direct Submission
JOURNAL      Submitted (15-MAY-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
On Mar 6, 2004 this sequence version replaced gi:32189675.
----- Genome Center
Center: Washington University Genome Sequencing Center
```


Center code: WUGSC
Web site: <http://genome.wustl.edu>
Contact: submissions@watson.wustl.edu
----- Summary Statistics -----
Center project name: M_BB0129H18

NOTICE:

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone, fosmid clone or direct clone walk sequence. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to obtain the consensus sequence; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-24 BAC library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

Location/Qualifiers

1. .207182

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="18"

/map="18"

/clone="RP24-129H18"

/clone_lib="RPCI-24"

17051. .17219

/note="Unresolved simple sequence repeat."

51528. .52399

/note="CpG island (%GC=71.6, o/e=0.81, #CpGs=79) "

105737. .105746

/note="Sequence derived from PCR product of project DNA."

128471. .128900

/note="Sequence derived from PCR product of project DNA."

146793. .147510

/note="Unresolved simple sequence repeat."

ORIGIN

Query Match 49.3%; Score 29.6; DB 10; Length 207182;

Best Local Similarity 73.1%; Pred. No. 43;

Matches 38; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 8 CAGCTCCTCCGCTCATGGGTACTTCTGCTCCAGATACTGCCCAGCTCC 59

DB 172796 CAGCTCCTCCCACTCCTCTGCTCCCTGCTCCTCCACTCCCCAGCTCC 172847

RESULT 15

AC124587

LOCUS AC124587 219963 bp DNA linear HTG 10-JUL-2004
DEFINITION Mus musculus chromosome 1 clone RP23-132G24, WORKING DRAFT
SEQUENCE, 5 unordered pieces.

ACCESSION AC124587

VERSION AC124587.4 GI:50201932

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu>
Contact: submissions@watson.wustl.edu
----- Project Information -----
Center project name: M_BA0132G24
----- Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing vector: plasmid; 100%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 217311 bases at least Q40
Consensus quality: 217860 bases at least Q30
Consensus quality: 218107 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 287: contig of 287 bp in length
* 288 387: gap of unknown length
* 388 9268: contig of 8881 bp in length
* 9269 9368: gap of unknown length
* 9369 36735: contig of 27367 bp in length
* 36736 36835: gap of unknown length
* 36836 76965: contig of 40130 bp in length
* 76966 77065: gap of unknown length
* 77066 219963: contig of 142898 bp in length.

Location/Qualifiers
1. .219963
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="1"
/clone="RP23-132G24"
1. .287
/note="assembly_name:Contig36
clone_end:T7
vector_side:right"
388. .9268
/note="assembly_name:Contig38"
9369. .36735
/note="assembly_name:Contig39"
36836. .76965
/note="assembly_name:Contig40"
77066. .219963
/note="assembly_name:Contig41
clone_end:SP6

FEATURES

source

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 22:37:00 ; Search time 183.4 Seconds
(without alignments)
1879.789 Million cell updates/sec

Title: US-10-057-136-5

Perfect score: 60
Sequence: 1 ggatccaccgcgcgcgcctgc.....ccgacacgcgcgcctccc 60

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10D_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	14 US-10-057-136-5	Sequence 5, Appli
2	60	100.0	2297	16 US-10-406-317-41	Sequence 41, Appl
3	38.2	63.7	60	14 US-10-057-136-11	Sequence 11, Appl
4	35	58.3	60	14 US-10-057-136-2	Sequence 2, Appli
5	35	58.3	1424	16 US-10-447-839A-75	Sequence 75, Appl
6	35	58.3	1428	16 US-10-447-839A-20	Sequence 20, Appl
7	35	58.3	1527	14 US-10-057-136-19	Sequence 19, Appl
8	35	58.3	1721	9 US-09-864-864-280	Sequence 280, App
9	35	58.3	1721	9 US-09-967-768A-224	Sequence 224, App
10	35	58.3	1721	14 US-10-247-703-21	Sequence 21, Appl
11	35	58.3	1721	14 US-10-097-340-211	Sequence 211, App
12	35	58.3	1721	14 US-10-171-311-155	Sequence 155, App

13	35	58.3	1721	15 US-10-007-926A-58	Sequence 58, Appl
14	35	58.3	1721	15 US-10-029-517-3	Sequence 3, Appli
15	35	58.3	1721	15 US-10-172-118-775	Sequence 775, App
16	35	58.3	1721	16 US-10-342-887-775	Sequence 775, App
17	35	58.3	1721	17 US-10-775-920-88	Sequence 88, Appl
18	35	58.3	1799	16 US-10-447-839A-19	Sequence 19, Appl
19	35	58.3	1804	9 US-09-964-824A-573	Sequence 573, App
20	35	58.3	1804	15 US-10-029-517-17	Sequence 17, Appl
21	35	58.3	1804	17 US-10-717-597-30	Sequence 30, Appl
22	35	58.3	1804	17 US-10-775-920-84	Sequence 84, Appl
23	35	58.3	1823	15 US-10-101-510-339	Sequence 339, App
24	35	58.3	2026	14 US-10-198-846-12589	Sequence 12589, A
25	35	58.3	2238	17 US-10-775-920-87	Sequence 87, Appl
26	35	58.3	2678	15 US-10-252-157-103	Sequence 103, App
27	35	58.3	4139	9 US-09-964-824A-105	Sequence 105, App
28	35	58.3	4139	9 US-09-864-824A-578	Sequence 578, App
29	35	58.3	4139	9 US-09-864-864-334	Sequence 334, App
30	35	58.3	4139	9 US-09-880-107-2121	Sequence 2121, Ap
31	35	58.3	4139	11 US-09-968-007A-751	Sequence 751, App
32	35	58.3	4139	14 US-10-171-311-157	Sequence 157, App
33	35	58.3	4139	15 US-10-177-293-310	Sequence 310, App
34	35	58.3	4139	16 US-10-440-464-155	Sequence 155, App
35	35	58.3	4139	17 US-10-734-564-53	Sequence 53, Appl
36	35	58.3	4139	17 US-10-775-920-80	Sequence 80, Appl
37	35	58.3	4139	17 US-10-775-920-85	Sequence 85, Appl
38	35	58.3	8181	16 US-10-447-839A-18	Sequence 18, Appl
39	35	58.3	8186	14 US-10-247-703-23	Sequence 23, Appl
40	35	58.3	8186	15 US-10-029-517-19	Sequence 19, Appl
41	33.6	56.0	518	14 US-10-247-703-38	Sequence 38, Appl
42	33.6	56.0	518	15 US-10-029-517-101	Sequence 101, App
43	33.6	56.0	3343	14 US-10-247-703-27	Sequence 27, Appl
44	33.6	56.0	3343	15 US-10-029-517-102	Sequence 102, Appl
45	33.4	55.7	78	14 US-10-057-136-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-10-057-136-5
; Sequence 5, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOW, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALLI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: TUMOR-ASSOCIATED ANTIGEN
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; PRIOR APPLICATION NUMBER: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-5

Query Match 100.0%; Score 60; DB 14; Length 60;
Best Local Similarity 100.0%; Pred. No. 4.3e-11;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCTGCGACGAGTGACGTGCGCGCCGACACGCGCCGCTCCC 60
|||||

Db 1 GGATCCACCGCGCGCTGCGCACGAGTGACGTGCGCGCCCGACACGCGCCCGCTCCC 60

RESULT 2

US-10-406-317-41
; Sequence 41, Application US/10406317
; Publication No. US20040019195A1
; GENERAL INFORMATION:
; APPLICANT: Schlom, Jeffrey;
; APPLICANT: Hodge, James;
; APPLICANT: Panicali, Dennis;
; TITLE OF INVENTION: A recombinant vector expressing multiple constitutulatory
; FILE REFERENCE: 38163-0189
; CURRENT APPLICATION NUMBER: US/10/406,317
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US/09/856,988
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: PCT/US99/26866
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/111,582
; PRIOR FILING DATE: 1998-12-09
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: VECTOR
US-10-406-317-41

Query Match 100.0%; Score 60; DB 16; Length 2297;
Best Local Similarity 100.0%; Pred. No. 1.9e-11;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCTGCGCACGAGTGACGTGCGCGCCCGACACGCGCCCGCTCCC 60
Db 466 GGATCCACCGCGCGCTGCGCACGAGTGACGTGCGCGCCCGACACGCGCCCGCTCCC 525

RESULT 3

US-10-057-136-11
; Sequence 11, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUEF, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-11

Query Match 63.7%; Score 38.2; DB 14; Length 60;
Best Local Similarity 78.0%; Pred. No. 0.001;

Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCTGCGCACGAGTGACGTGCGCGCCCGACACGCGCCCGCTCC 59
Db 1 GGTTCACGCGACCTCCAGACACGAGTGACGTGCACTGCACTCCGACACCGCTCAGCTCC 59

RESULT 4

US-10-057-136-2
; Sequence 2, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUEF, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-2

Query Match 58.3%; Score 35; DB 14; Length 60;
Best Local Similarity 74.6%; Pred. No. 0.013;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCTGCGCACGAGTGACGTGCGCGCCCGACACGCGCCCGCTCC 59
Db 1 GGCTCCACCGCGCGCCCGCCAGCCAGCGGTGTACCTGCGCGCCCGGACACGAGCGCGCC 59

RESULT 5

US-10-447-839A-75/c
; Sequence 75, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald W.
; APPLICANT: Kharbanda, Suresh
; APPLICANT: Weltman, Steven D.
; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT APPLICATION NUMBER: US/10/447,839A
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 75
; LENGTH: 1424
; TYPE: RNA
; ORGANISM: ARTIFICIAL
; FEATURE:
; OTHER INFORMATION: Synthesized Sequence
US-10-447-839A-75

Query Match 58.3%; Score 35; DB 16; Length 1424;
Best Local Similarity 74.6%; Pred. No. 0.0061;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCCGCTGCGCAGGAGTGACGTGCGCGCCCGACACGCGCCCGCTCC 59
|||
Db 1041 GGCTCCACCGCCCCCGCCAGCCACGCGTGTCACCTCGCGCCCGACACGCGCGCCCG 983

RESULT 6

US-10-447-839A-20
; Sequence 20, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald W.
; APPLICANT: Kharbanda, Surender
; APPLICANT: Weitman, Steven D.
; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT APPLICATION NUMBER: US/10/447,839A
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 1428
; TYPE: RNA
; ORGANISM: RNA
US-10-447-839A-20

Query Match 58.3%; Score 35; DB 16; Length 1428;
Best Local Similarity 69.5%; Pred. No. 0.0061;
Matches 41; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCCGCTGCGCAGGAGTGACGTGCGCGCCCGACACGCGCCCGCTCC 59
|||
Db 385 GGCUCACCGCGCCCGCCAGCCACGCGUGACCUCCGCCCGGACACGCGCGCGCC 443

RESULT 7

US-10-057-136-19
; Sequence 19, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: Kufe, DONALD
; APPLICANT: PANTCALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 1527
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-19

Query Match 58.3%; Score 35; DB 14; Length 1527;
Best Local Similarity 74.6%; Pred. No. 0.0061;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCCGCTGCGCAGGAGTGACGTGCGCGCCCGACACGCGCCCGCTCC 59
|||
Db 226 GGCTCCACCGCCCCCGCCAGCCACGCGTGTCACCTCGCGCCCGACACGCGCGCCCG 284

RESULT 8

US-09-864-864-280
; Sequence 280, Application US/09864864
; Patent No. US20020102679A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Secrist, Heather
; APPLICANT: Lodes, Michael J.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steve P.
; APPLICANT: Mannion, Jane
; APPLICANT: Benson, Darin R.
; APPLICANT: Carter, Darlick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.523
; CURRENT APPLICATION NUMBER: US/09/864,864
; CURRENT FILING DATE: 2001-05-23
; NUMBER OF SEQ ID NOS: 341
; SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 280
; LENGTH: 1721
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-864-864-280

Query Match 58.3%; Score 35; DB 9; Length 1721;
Best Local Similarity 74.6%; Pred. No. 0.0059;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCCGCTGCGCAGGAGTGACGTGCGCGCCCGACACGCGCCCGCTCC 59
|||
Db 562 GGCTCCACCGCGCCCGCCAGCCACGCGTGTCACCTCGCGCCCGGACACGCGCGCCCG 620

RESULT 9

US-09-967-768A-224
; Sequence 224, Application US/09967768A
; Patent No. US20020150877A1
; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-72
; CURRENT APPLICATION NUMBER: US/09/967,768A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/60/236,109
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,034
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,111
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 325
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 224
; LENGTH: 1721
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-967-768A-224

Query Match 58.3%; Score 35; DB 9; Length 1721;

Best Local Similarity 74.6%; Pred. No. 0.0059;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCCGCTGCGCAGCAGTGAAGTGCCTGGCGCCCGACACAGCGCCCGCTCC 59
|||
Db 562 GGCTCCACCGCGCCCGCAGCCCAAGGTGTCACTCGGCCCGGACACCAAGCGCGCC 620

RESULT 10
US-10-247-703-21

; Sequence 21, Application US/10247703
; Publication No. US20030063597A1
; GENERAL INFORMATION:
; APPLICANT: Branigan, Patrick
; APPLICANT: Goletz, Theresa J
; APPLICANT: Knight, David M
; APPLICANT: McCarthy, Stephen G
; APPLICANT: Scallion, Bernard J
; APPLICANT: Snyder, Linda A
; TITLE OF INVENTION: CYTOKINE ADJUVANT ENCODING NUCLEIC ACIDS
; FILE REFERENCE: CEN310
; CURRENT APPLICATION NUMBER: US/10/247,703
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: 60/328,371
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 1721
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-247-703-21

Query Match 58.3%; Score 35; DB 14; Length 1721;
Best Local Similarity 74.6%; Pred. No. 0.0059;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCCGCTGCGCAGCAGTGAAGTGCCTGGCGCCCGACACAGCGCCCGCTCC 59
|||
Db 562 GGCTCCACCGCGCCCGCAGCCCAAGGTGTCACTCGGCCCGGACACCAAGCGCGCC 620

RESULT 11
US-10-097-340-211

; Sequence 211, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26

; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 211
; LENGTH: 1721
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-097-340-211

Query Match 58.3%; Score 35; DB 14; Length 1721;
Best Local Similarity 74.6%; Pred. No. 0.0059;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCCGCTGCGCAGCAGTGAAGTGCCTGGCGCCCGACACAGCGCCCGCTCC 59
|||
Db 562 GGCTCCACCGCGCCCGCAGCCCAAGGTGTCACTCGGCCCGGACACCAAGCGCGCC 620

RESULT 12
US-10-171-311-155

; Sequence 155, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersch, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 155
; LENGTH: 1721
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-171-311-155

Query Match 58.3%; Score 35; DB 14; Length 1721;
Best Local Similarity 74.6%; Pred. No. 0.0059;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCCGCTGCGCAGCAGTGAAGTGCCTGGCGCCCGACACAGCGCCCGCTCC 59
|||
Db 562 GGCTCCACCGCGCCCGCAGCCCAAGGTGTCACTCGGCCCGGACACCAAGCGCGCC 620

RESULT 13

US-10-007-926A-58
; Sequence 58, Application US/10007926A
; Publication No. US20030143539A1

```
; GENERAL INFORMATION:
; APPLICANT: BERTUCCI, FRANCOIS
; APPLICANT: HOUUGATTE, REMI
; APPLICANT: BIRNBAUM, DANIEL
; APPLICANT: NGUYEN, CATHERINE
; APPLICANT: VIENS, PATRICE
; APPLICANT: FERT, VINCENT
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF PRIMARY BREAST CARCINOMAS
; TITLE OF INVENTION: USING ARRAYS OF CANDIDATE GENES
; FILE REFERENCE: 1546-R-00
; CURRENT APPLICATION NUMBER: US/10/007, 926A
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 60/254, 090
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 468
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 58
; LENGTH: 1721
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: mucin 1, transmembrane (MUC1) gene.
US-10-007-926A-58

Query Match          58.3%; Score 35; DB 15; Length 1721;
Best Local Similarity 74.6%; Pred. No. 0.0059;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY      1 GGATCCACCGCGCGCGCTGCGCAGGAGTGACGTGCGCGCCGACACGCGCCCGCTCC 59
DB      562 GGCTCCACCGCGCGCGCAGCCACGCGTGTACCTCGCGCCCGACACGAGCGCGCC 620
```

```
RESULT 14
US-10-029-517-3
; Sequence 3, Application US/10029517
; Publication No. US20030148969A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029, 517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 3
; LENGTH: 1721
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (58) ... (1605)
US-10-029-517-3
```

```
Query Match          58.3%; Score 35; DB 15; Length 1721;
Best Local Similarity 74.6%; Pred. No. 0.0059;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY      1 GGATCCACCGCGCGCGCTGCGCAGGAGTGACGTGCGCGCCGACACGCGCCCGCTCC 59
DB      562 GGCTCCACCGCGCGCGCAGCCACGCGTGTACCTCGCGCCCGACACGAGCGCGCC 620
```

```
RESULT 15
US-10-172-118-775
; Sequence 775, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
```

```
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172, 118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380, 770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2639
; SEQ ID NO 775
; LENGTH: 1721
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM 002456
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-775
```

```
Query Match          58.3%; Score 35; DB 15; Length 1721;
Best Local Similarity 74.6%; Pred. No. 0.0059;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY      1 GGATCCACCGCGCGCGCTGCGCAGGAGTGACGTGCGCGCCGACACGCGCCCGCTCC 59
DB      562 GGCTCCACCGCGCGCGCAGCCACGCGTGTACCTCGCGCCCGACACGAGCGCGCC 620
```

Search completed: January 16, 2005, 09:30:24
Job time : 184.4 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using BW model

Run on: January 15, 2005, 19:50:13 ; Search time 35.2 Seconds
(without alignments)
1211.572 Million cell updates/sec

Title: US-10-057-136-5

Perfect score: 60

Sequence: 1 ggatccaccgcgcgcctctgc.....ccgacacgcgcgcctccc 60

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued Patents NA: *
2: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/PCtUS_COMB.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	58.3	1721	4 US-10-029-517-3	Sequence 3, Appli
2	35	58.3	1804	4 US-10-029-517-17	Sequence 17, Appli
3	35	58.3	8186	4 US-10-029-517-19	Sequence 19, Appli
4	33.6	56.0	518	4 US-10-029-517-101	Sequence 101, App
5	33.6	56.0	3343	4 US-10-029-517-102.	Sequence 102, App
6	33.4	55.7	572	4 US-10-029-517-18	Sequence 18, Appli
7	32.2	53.7	60	4 US-09-475-947A-246	Sequence 246, App
8	32	53.3	981	4 US-10-029-517-16	Sequence 16, Appli
9	30.8	51.3	6192	2 US-08-479-537A-1	Sequence 1, Appli
10	30.8	51.3	6192	3 US-09-083-116-1	Sequence 1, Appli
11	30.8	51.3	6192	3 US-09-134-916A-1	Sequence 1, Appli
12	30.8	51.3	6449	2 US-08-479-537A-4	Sequence 4, Appli
13	30.8	51.3	6449	3 US-09-083-116-4	Sequence 4, Appli
14	30.8	51.3	6449	3 US-09-134-916A-4	Sequence 4, Appli
15	29.2	48.7	1023	1 US-08-252-966B-16	Sequence 16, Appli
16	27.8	46.3	2574	4 US-09-780-045-3	Sequence 3, Appli
17	27.2	45.3	1942	4 US-10-140-002-515	Sequence 515, App
18	26.4	44.0	1026	4 US-09-489-039A-1454	Sequence 1454, Ap
19	26.2	43.7	759	4 US-09-252-991A-5901	Sequence 5901, Ap
20	26.2	43.7	900	4 US-09-252-991A-5969	Sequence 5969, Ap
21	26.2	43.7	954	4 US-09-252-991A-5918	Sequence 5918, Ap
22	25.4	42.3	4403765	3 US-09-103-840A-2	Sequence 2, Appli
23	25.4	42.3	4403765	3 US-09-103-840A-2	Sequence 2, Appli
24	25.4	42.3	4411529	3 US-09-103-840A-1	Sequence 1, Appli
25	25.4	42.3	4411529	3 US-09-103-840A-1	Sequence 1, Appli
26	25.2	42.0	1029	4 US-09-252-991A-15460	Sequence 15460, A
27	25.2	42.0	1098	4 US-09-252-991A-15400	Sequence 15400, A

C	28	25.2	42.0	1488	4 US-09-252-991A-15378	Sequence 15378, A
	29	25.2	42.0	1842	4 US-09-252-991A-15508	Sequence 15508, A
	30	25.2	42.0	2877	3 US-09-235-103-1	Sequence 1, Appli
	31	25	41.7	1824	4 US-09-799-451-301	Sequence 301, App
	32	24.8	41.3	357	4 US-09-252-991A-441	Sequence 441, App
C	33	24.8	41.3	1445	4 US-09-614-912-77	Sequence 77, Appli
	34	24.8	41.3	1446	4 US-09-252-991A-384	Sequence 384, App
	35	24.8	41.3	1488	4 US-09-252-991A-421	Sequence 421, App
	36	24.8	41.3	2825	3 US-09-196-390-5	Sequence 5, Appli
	37	24.8	41.3	2825	4 US-09-952-677-5	Sequence 5, Appli
	38	24.8	41.3	5993	3 US-09-383-630-1	Sequence 1, Appli
	39	24.8	41.3	5993	3 US-08-804-227C-1	Sequence 1, Appli
	40	24.8	41.3	43280	2 US-08-403-852D-9	Sequence 9, Appli
	41	24.6	41.0	645	2 US-08-510-646B-9	Sequence 9, Appli
	42	24.6	41.0	645	3 US-09-231-818-9	Sequence 9, Appli
	43	24.6	41.0	645	3 US-09-635-359B-9	Sequence 9, Appli
	44	24.6	41.0	645	4 US-08-765-907A-14	Sequence 14, Appli
	45	24.6	41.0	2220	3 US-08-765-907A-14	Sequence 14, Appli

ALIGNMENTS

RESULT 1
US-10-029-517-3
; Sequence 3, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 3
; LENGTH: 1721
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (58)...(1605)
US-10-029-517-3

Query Match 58.3%; Score 35; DB 4; Length 1721;
Best Local Similarity 74.6%; Pred. No. 0.041;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCTTGGACGAGTGACGTGGCGCCGACACGCGCCGCTCC 59
Db 562 GGCTCCACCGCGCGCCGACGAGTGTACCTCGCGCCCGACACGCGCGCC 620

RESULT 2
US-10-029-517-17
; Sequence 17, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 17
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (73)...(1500)
US-10-029-517-17

Query Match 58.3%; Score 35; DB 4; Length 1804;
Best Local Similarity 74.6%; Pred. No. 0.041;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCCCTGGCGCAGAGTGACGTGCGCGCCGACACAGCGCCCGCTCC 59
Db 457 GGCTCCACCGCGCGCCCGCCAGCCCAAGGTGTCACCTCGCGCCCGACACAGCGCGCCCG 515

RESULT 3

US-10-029-517-19
; Sequence 19, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 19
; LENGTH: 8186
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 6899
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7155
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7184
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7957
; OTHER INFORMATION: unknown
; NAME/KEY: intron
; LOCATION: (2997)...(3498)
; OTHER INFORMATION: intron 1
; NAME/KEY: intron:exon junction
; LOCATION: (3498)...(3499)
; OTHER INFORMATION: intron 1:exon 2
; NAME/KEY: exon
; LOCATION: (3508)...(3599)
; OTHER INFORMATION: exon 2d
; NAME/KEY: exon:intron junction
; LOCATION: (3982)...(3983)
; OTHER INFORMATION: exon 2a:intron 2a
; NAME/KEY: intron:exon junction
; LOCATION: (4205)...(4206)
; OTHER INFORMATION: intron 2c:exon 3c
; NAME/KEY: intron:exon junction
; LOCATION: (4259)...(4260)
; OTHER INFORMATION: intron 2d:exon 3d
; NAME/KEY: exon
; LOCATION: (4260)...(4328)
; OTHER INFORMATION: exon 3d
; NAME/KEY: intron:exon junction
; LOCATION: (4632)...(4633)
; OTHER INFORMATION: intron 3:exon 4
; NAME/KEY: exon
; LOCATION: (4914)...(5035)
; OTHER INFORMATION: exon 5
; NAME/KEY: intron
; LOCATION: (5266)...(6293)
; OTHER INFORMATION: intron 6
US-10-029-517-19

Query Match 58.3%; Score 35; DB 4; Length 8186;
Best Local Similarity 74.6%; Pred. No. 0.039;

Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCCCTGGCGCAGAGTGACGTGCGCGCCGACACAGCGCCCGCTCC 59
Db 3825 GGCTCCACCGCGCGCCCGCCAGCCCAAGGTGTCACCTCGCGCCCGACACAGCGCGCCCG 3883

RESULT 4

US-10-029-517-101
; Sequence 101, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 101
; LENGTH: 518
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-10-029-517-101

Query Match 56.0%; Score 33.6; DB 4; Length 518;
Best Local Similarity 75.0%; Pred. No. 0.11;
Matches 42; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCCCTGGCGCAGAGTGACGTGCGCGCCGACACAGCGCCCGCGC 56
Db 462 GGCTCCACCGCGCGCCCGCCAGCCCAAGGTGTCACCTCGCGCCCGACACAGCGCGCGC 517

RESULT 5

US-10-029-517-102
; Sequence 102, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 102
; LENGTH: 3343
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-10-029-517-102

Query Match 56.0%; Score 33.6; DB 4; Length 3343;
Best Local Similarity 75.0%; Pred. No. 0.1;
Matches 42; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCCCTGGCGCAGAGTGACGTGCGCGCCGACACAGCGCCCGCGC 56
Db 1728 GGCTCCACCGCGCGCCCGCCAGCCCAAGGTGTCACCTCGCGCCCGACACAGCGCGCGC 1783

RESULT 6

US-10-029-517-18
; Sequence 18, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517

; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 18
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (67)...(572)
US-10-029-517-18

Query Match 55.7%; Score 33.4; DB 4; Length 572;
Best Local Similarity 72.9%; Pred. No. 0.13;
Matches 43; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

OY 1 GGATCCACCGCGCGCTGGCGACGAGTGAGTCGGCGCCCGACACGCGCCCGCTCC 59
|||
Db 478 GGCTCCACCGCGCCCGCCCAAGCCACGCGTGTCACTCGGCGCCCGGACACGAGCGCGCCCG 536

RESULT 7

US-09-475-947A-246
; Sequence 246, Application US/09475947A
; Patent No. 6472154
; GENERAL INFORMATION:
; APPLICANT: Garner, Harold R.
; APPLICANT: Wren, Jonathan D.
; APPLICANT: Minna, John D.
; TITLE OF INVENTION: Polymorphic Repeats in Human Genes
; FILE REFERENCE: UTS0667
; CURRENT APPLICATION NUMBER: US/09/475, 947A
; CURRENT FILING DATE: 1999-12-31
; NUMBER OF SEQ ID NOS: 346
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 246
; LENGTH: 60
; TYPE: DNA
; ORGANISM: human
US-09-475-947A-246

Query Match 53.7%; Score 32.2; DB 4; Length 60;
Best Local Similarity 75.5%; Pred. No. 0.3;
Matches 40; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

OY 1 GGATCCACCGCGCGCTGGCGACGAGTGAGTCGGCGCCCGACACGCGCCCG 53
|||
Db 7 GGCTCCACCGCGCCCGCCCAAGCCACGCGTGTCACTCGGCGCCCGGACACGAGCGCC 59

RESULT 8

US-10-029-517-16
; Sequence 16, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029, 517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 16
; LENGTH: 981
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon:exon junction
; LOCATION: (464)...(465)
; OTHER INFORMATION: exon 3b:exon 4
US-10-029-517-16

Query Match 53.3%; Score 32; DB 4; Length 981;
Best Local Similarity 73.2%; Pred. No. 0.32;

Matches 41; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
OY 1 GGATCCACCGCGCGCTGGCGACGAGTGAGTCGGCGCCCGACACGCGCCCGCC 56
|||
Db 21 GGCTCCACCGCGCCCGCCCAAGCCACGCGTGTCACTCGGCGCCCGGACACAGGCCCGC 76

RESULT 9

US-08-479-537A-1
; Sequence 1, Application US/08479537A
; Patent No. 5861381
; GENERAL INFORMATION:
; APPLICANT: CHAMON, Pierre
; APPLICANT: KIENY, Marie-Paule
; APPLICANT: LATHE, Richard
; APPLICANT: HAREUVENI, Mara
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
; TITLE OF INVENTION: TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479, 537A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 90/13101
; FILING DATE: 23-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR91/00835
; FILING DATE: 23-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/039,320
; FILING DATE: 04-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,576
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teekin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 017753-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6192 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 58..120
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 439..5239
; OTHER INFORMATION: /note= "The nucleotides spanning
; OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
; OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
; OTHER INFORMATION: The number of such repeats varies from 1 to 80."
; FEATURE:
; NAME/KEY: mat_peptide

```
LOCATION: 121..6166
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-08-479-537A-1
```

```
Query Match
Best Local Similarity 51.3%; Score 30.8; DB 2; Length 6192;
Matches 38; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
```

```
OY 1 GGATCCACCGCGCGCTGCGCAGGAGTGACGTGCGCGCCGACACGCGCCGCTCC 59
Db 442 GGCTCCACCGCGCGCCNNNGCCACGCGTGTCACCTCGGCGCCGACGACNNAGCGCCGNNCC 500
```

RESULT 10

```
US-09-083-116-1
Sequence 1, Application US/09083116
Patent No. 6203795
```

```
GENERAL INFORMATION:
APPLICANT: CHAMBERN, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREUVENI, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
```

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/083,116
FILING DATE:
```

```
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,537
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
```

```
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
```

```
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6192 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
```

```
MOLECULE TYPE: DNA (genomic)
```

```
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
```

```
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
```

```
OTHER INFORMATION: /note= "The nucleotides spanning
OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
OTHER INFORMATION: The number of such repeats varies from 1 to 80."
```

```
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 121..6166
```

```
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
```

```
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
```

```
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
```

```
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
```

```
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
```

```
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
```

```
US-09-083-116-1
```

```
Query Match
Best Local Similarity 51.3%; Score 30.8; DB 3; Length 6192;
Matches 38; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
```

```
OY 1 GGATCCACCGCGCGCTGCGCAGGAGTGACGTGCGCGCCGACACGCGCCGCTCC 59
Db 442 GGCTCCACCGCGCGCCNNNGCCACGCGTGTCACCTCGGCGCCGACGACNNAGCGCCGNNCC 500
```

RESULT 11

```
US-09-134-916A-1
Sequence 1, Application US/09134916A
Patent No. 6328956
```

```
GENERAL INFORMATION:
APPLICANT: CHAMBERN, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREUVENI, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
```

```
COMPUTER READABLE FORM:
```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/134,916A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/479,537
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: FR 90/13101
FILING DATE: 23-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teekin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6192 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note="The nucleotides spanning
OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
OTHER INFORMATION: The number of such repeats varies from 1 to 80."
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 121..6166
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note="Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note="Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note="Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-09-134-916A-1
Query Match 51.3%; Score 30.8; DB 3; Length 6192;
Best Local Similarity 64.4%; Pred. No. 0.67;
Matches 38; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCCCTGCGCAGGAGTGACGTCCGCGCCCGACACGCGCCCGCTCC 59
Db 442 GGCTCCACCGCGCCCGCCNNNGCCACCGGTGTCACTCGCGCCCGGACNNNAGGCCGNNCC 500
RESULT 12
US-08-479-537A-4
Sequence 4, Application US/08479537A
Patent No. 5861381
GENERAL INFORMATION:
APPLICANT: CHAMON, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREUVENI, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/479,537A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 90/13101
FILING DATE: 23-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teekin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6449 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note="The nucleotides spanning
OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
OTHER INFORMATION: The number of such repeats varies from 1 to 80."
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 121..5661
FEATURE:

```
; NAME/KEY: repeat_region
; LOCATION: 457
; OTHER INFORMATION: /note="Nucleotide 457 is X1 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 487
; OTHER INFORMATION: /note="Nucleotide 487 is Y = NNN
; OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
; OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 496
; OTHER INFORMATION: /note="Nucleotide 496 is X2 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; US-08-479-537A-4

Query Match      51.3%; Score 30.8; DB 2; Length 6449;
Best Local Similarity 64.4%; Pred. No. 0.67;
Matches 38; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY      1  GGATCCACCGCGCGCCTGCGCAGGAGTGACGTGCGCGCCGACACGCGCCGCTCC 59
Db      442 GGCTCCACCGCGCGCCNNNGCCACGCGTGTCACCTCGCGCCCGGACNNNAGGCCGNNCC 500

RESULT 13
US-09-083-116-4
; Sequence 4, Application US/09083116
; GENERAL INFORMATION:
; APPLICANT: CHAMBON, Pierre
; APPLICANT: KIENY, Marie-Paule
; APPLICANT: LATHE, Richard
; APPLICANT: HAREUVENT, Mara
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
; TITLE OF INVENTION: TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/083,116
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,537
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR91/00835
; FILING DATE: 23-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/039,320
; FILING DATE: 04-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,576
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 017753-025
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6449 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 58..120
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 439..5239
; OTHER INFORMATION: /note="The nucleotides spanning
; OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
; OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
; OTHER INFORMATION: The number of such repeats varies from 1 to 80."
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 121..5661
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 457
; OTHER INFORMATION: /note="Nucleotide 457 is X1 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 487
; OTHER INFORMATION: /note="Nucleotide 487 is Y = NNN
; OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
; OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 496
; OTHER INFORMATION: /note="Nucleotide 496 is X2 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; US-09-083-116-4

Query Match      51.3%; Score 30.8; DB 3; Length 6449;
Best Local Similarity 64.4%; Pred. No. 0.67;
Matches 38; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY      1  GGATCCACCGCGCGCCTGCGCAGGAGTGACGTGCGCGCCGACACGCGCCGCTCC 59
Db      442 GGCTCCACCGCGCGCCNNNGCCACGCGTGTCACCTCGCGCCCGGACNNNAGGCCGNNCC 500

RESULT 14
US-09-134-916A-4
; Sequence 4, Application US/09134916A
; Patent No. 6328956
; GENERAL INFORMATION:
; APPLICANT: CHAMBON, Pierre
; APPLICANT: KIENY, Marie-Paule
; APPLICANT: LATHE, Richard
; APPLICANT: HAREUVENT, Mara
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
; TITLE OF INVENTION: TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```



```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/134,916A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,537
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: FR 90/13101
; FILING DATE: 23-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR91/00835
; FILING DATE: 23-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/039,320
; FILING DATE: 04-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,576
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 017753-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6449 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 58..120
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 439..5239
; OTHER INFORMATION: /note= "The nucleotides spanning
; OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
; OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
; OTHER INFORMATION: The number of such repeats varies from 1 to 80."
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 121..5661
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 457
; OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 487
; OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
; OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
; OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 496
; OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; US-09-134-916A-4

```

```

Query Match      51.3%; Score 30.8; DB 3; Length 6449;
Best Local Similarity 64.4%; Pred. No. 0.67;
Matches 38; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```

```

QY 1 GGATCCACCGCGCGCGCTGCGACGAGTGACGTGCGCGCCCGACACGCGCGCTCC 59
    || || || || || || || || || || || || || || || || || || ||

```

```

Db 442 GGCTCCACCGCCCCNNNGCCACAGGTGTACCTCGCCCCCGACNNNAGCCGNNCC 500
; RESULT 15
; US-08-252-966B-16/C
; Sequence 16, Application US/08252966B
; Patent No. 5624818
; GENERAL INFORMATION:
; APPLICANT: Eisenman, Robert N.
; APPLICANT: Hurlin, Peter J.
; APPLICANT: Ayer, Donald E.
; TITLE OF INVENTION: Regulatory Proteins that Dimerize with
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Christensen, O'Connor, Johnson, and Kindness PLLC
; STREET: 1420 Fifth Ave., Suite 2800
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101-2347
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/252,966B
; FILING DATE: 01-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Shelton, Dennis K.
; REGISTRATION NUMBER: 26,997.
; REFERENCE/DOCKET NUMBER: FHCRI7694
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 682-8100
; TELEFAX: (206) 224-0779
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1023 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; DESCRIPTION: clone 20; see Figure 27
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Mus musculus
; US-08-252-966B-16

```

```

Query Match      48.7%; Score 29.2; DB 1; Length 1023;
Best Local Similarity 69.0%; Pred. No. 2.1;
Matches 40; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```

```

QY 2 GATCCACCGCGCGCGCTGCGACGAGTGACGTGCGCGCCCGACACGCGCGCTCC 59
    || || || || || || || || || || || || || || || || || || ||
Db 67 GCTCATCTCTGCGCGCGCGCGACGAGTGACGTGCGCGCGCGCTCC 10

```

```

Search completed: January 16, 2005, 03:01:26
Job time : 45.2 secs

```

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 19:43:55 ; Search time 1550.6 Seconds
(without alignments)
1410.024 Million cell updates/sec

Title: US-10-057-136-5

Perfect score: 60

Sequence: 1 ggaaccacgcgcgcgcctgc.....ccgacacgcgcgcgcctccc 60

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	35	58.3	330	1	AI925867 wo20d04.x
2	35	58.3	604	4	BM791359 K-EST0071
3	35	58.3	877	5	BU542454 AGENCOURT
4	35	58.3	959	6	CA489836 AGENCOURT
5	35	58.3	1113	5	BU148487 AGENCOURT
6	35	58.3	1130	5	BU542996 AGENCOURT
7	35	58.3	1234	5	BQ936898 AGENCOURT
8	35	58.3	1268	5	BQ943554 AGENCOURT
9	35	58.3	1343	5	BQ920055 AGENCOURT
10	35	58.3	1349	5	BU152566 AGENCOURT
11	35	58.3	1420	5	BU542790 AGENCOURT
12	35	57.0	648	6	CF135986 UT-HF-BNO
C 13	34.2	57.0	1308	4	BM922810 AGENCOURT
14	33.4	55.7	1262	5	BQ935496 AGENCOURT
15	33.4	55.7	1531	5	BU543309 AGENCOURT
C 16	32.6	54.3	566	4	BI544718 AGENCOURT
C 17	32.6	54.3	761	8	AZ184492 SP_1003_A
C 18	32.6	54.3	786	6	CB990596 AGENCOURT
19	32	53.3	981	4	BG774910 602649832
20	32	53.3	1536	5	BQ923149 AGENCOURT
21	31.8	53.0	773	5	BU438869 603208290
22	31	51.7	166	7	T27692 EST12384 Hu
23	31	51.7	475	6	CB120860 K-EST0168
24	31	51.7	669	6	CB122585 K-EST0170

25	30.6	51.0	1735	9	CL469825	SAIL_133
26	30.4	50.7	740	8	AQ872562	nbe0051D
C 27	30.2	50.3	832	4	BG310404	HVSMC001
28	30	50.0	1037	9	AG359456	Mus muscu
29	29.4	49.0	878	9	CL471392	SAIL_159
30	29.4	49.0	1307	9	CL486103	SAIL_427
31	29.2	48.7	361	9	CNS0165F	AL106317
C 32	29.2	48.7	374	5	BY022312	Drosophi
C 33	29.2	48.7	389	5	BY096177	BY096177
C 34	29.2	48.7	437	2	BE653652	UI-M-AH1-
C 35	29.2	48.7	445	6	CB744569	AMGNNUC:M
C 36	29.2	48.7	462	5	BP770004	BP770004
C 37	29.2	48.7	570	4	BG148094	uu77b09.y
C 38	29.2	48.7	579	2	BE287253	601096117
C 39	29.2	48.7	581	7	CK432388	UI-D-GC1-
C 40	29.2	48.7	704	4	BI145614	602910642
C 41	29.2	48.7	782	4	BI738042	603357131
C 42	29.2	48.7	847	4	BI655851	603284868
C 43	29.2	48.7	910	4	BI654898	603282849
C 44	29.2	48.7	973	5	BQ230880	AGENCOURT
C 45	29.2	48.7	1024	6	BY707826	BY707826

ALIGNMENTS

RESULT 1
AI925867/c
LOCUS
DEFINITION
AI925867
wo20d04.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2455879 3'
similar to contains element MSRI MSRI repetitive element ;, mRNA
sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AI925867
AI925867.1 GI:5661831
EST.

REFERENCE
AUTHORS
TITLE
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
1 (bases 1 to 330)
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www-bio.lnl.gov/bbrp/image/image.html
Insert length: 578 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 305.
Location/Qualifiers

FEATURES
source
1..330
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2455879"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Pan1"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: Sali;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.72 kb. Life Technologies catalog #:
11548-013"

ORIGIN

Query Match 58.3%; Score 35; DB 1; Length 330;
Best Local Similarity 74.6%; Pred. No. 2.7;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCCTGCGCAGGAGTGCCTGGCGCCCGACACAGCGCCCGCTCC 59
Db 316 GGCTCCACCGCGCCCGCCAGCCCGACGCTGTCACTCGGCCCGCGACACAGCGCCGCCCC 258

RESULT 2

LOCUS BM791359 604 bp mRNA linear EST 05-MAR-2002
DEFINITION K-EST0071342 S21SNUS20 Homo sapiens cDNA clone S21SNUS20-14-A06 5',
mRNA sequence.

ACCESSION BM791359
VERSION BM791359.1 GI:19139591

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS 1 (bases 1 to 604)

Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.

21C Frontlier Korean EST Project 2001

Unpublished (2002)

CONTACT: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 14 row: A column: 06

High quality sequence stop: 604.

Location/Qualifiers

1..604

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="S21SNUS20-14-A06"

/sex="F"

/tissue_type="Stomach"

/cell_type="Floating aggregates"

/cell_line="SNU-520"

/lab_host="T0p10F"

/clone_lib="S21SNUS20"

/note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
Site_2: NotI; The poly (A) + RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dT-selected mRNA by
priming with dT-tailed vector. The dT-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli T0p10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN

Query Match 58.3%; Score 35; DB 4; Length 604;
Best Local Similarity 74.6%; Pred. No. 2.5;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCCTGCGCAGGAGTGCCTGGCGCCCGACACAGCGCCCGCTCC 59
Db 28 GGCTCCACCGCAGCCCGCCAGCCCGACGCTGTCACTCGGCCCGGACACAGCGCCGCCCC 86

RESULT 3
BU542454

LOCUS BU542454 877 bp mRNA linear EST 13-SEP-2002
DEFINITION AGENCOURT 10322173 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574322
5', mRNA sequence.

ACCESSION BU542454

VERSION BU542454.1 GI:22852937

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 877)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov

Tissue Procurement: DCTD/DTP

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1CM2769 row: h column: 02

High quality sequence stop: 760.

Location/Qualifiers

1..877

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6574322"

/tissue_type="carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_40"

/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Library constructed by

Ling Hong in the laboratory of Gerald M. Rubin (University

of California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH_MGC Library."

ORIGIN

Query Match 58.3%; Score 35; DB 5; Length 877;
Best Local Similarity 74.6%; Pred. No. 2.4;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCCTGCGCAGGAGTGCCTGGCGCCCGACACAGCGCCCGCTCC 59
Db 35 GGCTCCACCGCGCCCGCCAGCCCGACGCTGTCACTCGGCCCGGACACAGCGCCGCCCC 93

RESULT 4

CA489836

DEFINITION AGENCOURT 10810668 MAPCL Homo sapiens cDNA clone IMAGE:6722324 5',
mRNA sequence.

ACCESSION CA489836

VERSION CA489836.1 GI:24952627

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 959)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov

Tissue Procurement: Kristi A. Egland, Ira Pastan

cDNA Library Preparation: Invitrogen Corp

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov

Plate: LLM14284 row: n column: 20
High quality sequence start: 31
High quality sequence stop: 446.

FEATURES
source
1..959
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6722324"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231,
hTERT-HME1, LNCaP"
/lab_host="EMDH10B"
/clone_lib="MAPL"
/note="Vector: pCMV-SPORT6; Site_1: EcoRV; Site_2: Not I;
Subtracted with brain, liver, lung, kidney and muscle.
Directionally cloned. Priming method: oligo-dT. Average
insert size: 1800 bp. Library amplification: 26,000 fold.
Kristi A. Egland, James J. Vincent, Robert Strausberg,
Bungkook Lee & Ira Pastan: Discovery of new breast
cancer genes encoding membrane and secreted proteins.
Manuscript submitted."

ORIGIN

Query Match 58.3%; Score 35; DB 6; Length 959;
Best Local Similarity 74.6%; Pred. No. 2.4;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GGATCCACCGCGCGCCTGGCGACGAGTGACGTGCGCGCCCGACACGCGCCGCTCC 59
|||
Db 349 GGCTCCACCGCGCCCGCCAGCCACGCGTGTCACTCGCGCCCGGACACGAGCGCGCCC 407

RESULT 5
LOCUS BU148487 1113 bp mRNA linear EST 03-SEP-2002
DEFINITION AGENCOURT_8670479 NIH_MGC_40 Homo sapiens CDNA clone IMAGE:6380642
5', mRNA sequence.
ACCESSION BU148487
VERSION BU148487.1 GI:22662019
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
1 (bases 1 to 1113)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM2569 row: j column: 03
High quality sequence stop: 235.

FEATURES
source
1..1113
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6380642"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"

/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACGAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 58.3%; Score 35; DB 5; Length 1113;
Best Local Similarity 74.6%; Pred. No. 2.4;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GGATCCACCGCGCGCCTGGCGACGAGTGACGTGCGCGCCCGACACGCGCCGCTCC 59
|||
Db 90 GGCTCCACCGCGCCCGCCAGCCACGCGTGTCACTCGCGCCCGGACACGAGCGCGCCC 148

RESULT 6
LOCUS BU542996 1130 bp mRNA linear EST 13-SEP-2002
DEFINITION AGENCOURT_10338707 NIH_MGC_40 Homo sapiens CDNA clone IMAGE:6574979
5', mRNA sequence.
ACCESSION BU542996
VERSION BU542996.1 GI:22853479
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
1 (bases 1 to 1130)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM2771 row: c column: 11
High quality sequence start: 27
High quality sequence stop: 246.

FEATURES
source
1..1130
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6574979"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACGAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 58.3%; Score 35; DB 5; Length 1130;
Best Local Similarity 74.6%; Pred. No. 2.4;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GGATCCACCGCGCGCCTGGCGACGAGTGACGTGCGCGCCCGACACGCGCCGCTCC 59
|||
Db 69 GGCTCCACCGCGCCCGCCAGCCACGCGTGTCACTCGCGCCCGGACACGAGCGCGCCC 127

RESULT 7
BQ936898
LOCUS
DEFINITION BQ936898 1234 bp mRNA linear EST 21-AUG-2002
AGENCOURT 8919110 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6383855
5', mRNA sequence.
BQ936898
ACCESSION BQ936898.1 GI:22352281
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1234)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM2577 row: 0 column: 24
High quality sequence stop: 245.
Location/Qualifiers
1. 1234
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6383855"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
source
1. 1234
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6383855"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

Query Match 58.3%; Score 35; DB 5; Length 1234;
Best Local Similarity 74.6%; Pred. No. 2.3;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GGATCCACCGCGCGCTGGCGACGAGTGACGTGGGCGCCGACACGGCCCGCTCC 59
|||
Db 90 GGCTCCACCGCGCGCCCGACGCGGTGCACCTCGGCGCCCGACACGAGCGCGCCC 148

RESULT 8
BQ943554
LOCUS
DEFINITION BQ943554 1268 bp mRNA linear EST 21-AUG-2002
AGENCOURT 8777815 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6384308
5', mRNA sequence.
BQ943554
ACCESSION BQ943554.1 GI:22359032
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1268)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM2579 row: 0 column: 21
High quality sequence stop: 177.
Location/Qualifiers
1. 1268
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6384308"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
source
1. 1268
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6384308"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

Query Match 58.3%; Score 35; DB 5; Length 1268;
Best Local Similarity 74.6%; Pred. No. 2.3;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GGATCCACCGCGCGCTGGCGACGAGTGACGTGGGCGCCGACACGGCCCGCTCC 59
|||
Db 90 GGCTCCACCGCGCGCCCGACGCGGTGCACCTCGGCGCCCGACACGAGCGCGCCC 148

RESULT 9
BQ920055
LOCUS
DEFINITION BQ920055 1343 bp mRNA linear EST 20-AUG-2002
AGENCOURT 10031674 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6481609
5', mRNA sequence.
BQ920055
ACCESSION BQ920055.1 GI:22334753
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1343)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM2663 row: 0 column: 02
High quality sequence start: 56
High quality sequence stop: 237.
Location/Qualifiers
1. 1343
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6481609"

TITLE	Journal	Medline	Pubmed	Comment
Normalization and subtraction: two approaches to facilitate gene discovery	Genome Res. 6 (9), 791-806 (1996)	97044477	8889548	Contact: Soares, MB Coordinated Laboratory for Computational Genomics University of Iowa 375 Newton Road, 4156 MEBF, Iowa City, IA 52242, USA Tel: 319 335 8250 Fax: 319 335 9565 Email: bento-soares@uiowa.edu Tissue Procurement: Louis Staudt
CDNA library preparation: Dr. M. Bento Soares, University of Iowa	CDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa	CDNA Sequencing by: Dr. M. Bento Soares, University of Iowa	Clone Distribution: Distribution information can be found at http://genome.uiowa.edu/distribution/humanfl.html	The following repetitive elements were found in this cDNA sequence: 101-147, >GC_rich#low_complexity Seq primer: PYX-5.
FEATURES	source	1.648	Location/Qualifiers	
	/organism="Homo sapiens"		/mol_type="mRNA"	
	/db_xref="taxon:9606"		/clone="IMAGE:3091050"	
	/tissue_type="lymph"		/cell_type="germinal center B cells"	
	/cell_line="MGC85"		/lab_host="DH10B (LTI)"	
	/clone_lib="NIH_MGC_50"		/note="Vector: pRT3-Pac; Site 1: NotI; Site 2: Eco RI; Constructed from size fractionated cytoplasmic mRNA (3.5-4.4kb). Directionally cloned. Cells provided by Louis M. Staudt, Ph.D. Library preparation by Maria de Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."	
ORIGIN	Query Match	57.0%;	Score 34.2;	DB 6; Length 648;
	Best Local Similarity	76.4%;	Pred. No. 4.3;	
	Matches 42;	Conservative 0;	Mismatches 13;	Indels 0; Gaps 0;
Qy	6	CACCGCGCGCCTGGCGACGAGTGAAGTCTGGGCGCCGACACGCGCCGCTCCC	60	
Db	147	CCCCGCGCGCCTGGCGCTGTGACAGAGCCCGCGCGCCGCTCGGCTCTACTCCC	93	
RESULT 13	BM922810/c	LOCUS	BM922810	1308 bp mRNA linear EST 12-MAR-2002
DEFINITION	AGENCOURT_6652626 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5755931	5', mRNA	Sequence.	
ACCESSION	BM922810	VERSION	BM922810.1	GI:19373189
KEYWORDS	EST.	SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	NIH-MGC http://mgc.nci.nih.gov/ .		National Institutes of Health, Mammalian Gene Collection (MGC)	
AUTHORS	Unpublished (1999)		Contact: Robert Strausberg, Ph.D.	
TITLE	Email: cgabbs-r@mail.nih.gov		Tissue Procurement: Life Technologies, Inc.	
JOURNAL	CDNA Library Preparation: Life Technologies, Inc.		CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	
COMMENT	DNA Sequencing by: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov			

Plate:	LLAM12795	row:	1	column:	12
High quality sequence stop: 331.					
Location/Qualifiers					
1. .1308					
/organism="Homo sapiens"					
/mol_type="mRNA"					
/db_xref="taxon:9606"					
/clone="IMAGE:5755931"					
/tissue_type="leukocyte"					
/lab_host="DH10B"					
/clone_lib="NIH_MGC_118"					
/note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH_MGC Library."					
this is a NIH_MGC Library."					
ORIGIN					
Query Match					
Best Local Similarity 57.0%; Score 34.2; DB 4; Length 1308;					
Matches 42; Conservative 0; Mismatches 13; Indels 0; Gaps 0;					
Db					
6 CACCGCGCGCGCTGCGCAGGAGTGACGTGCGCGCCGACACGCGCCGCTCCC 60					
80 CCCC GCCCGCGCTGCGCTGCGCTGAGAGCCCGCGCGCGCGCTGCGCTACTCCC 26					
RESULT 14					
BQ935496					
LOCUS					
DEFINITION					
BQ935496 1262 bp mRNA linear EST 21-AUG-2002					
AGENCOURT_8776475 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6382088					
5', mRNA sequence.					
BQ935496					
BQ935496.1 GI:22350879					
EST.					
Homo sapiens (human)					
Homo sapiens					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.					
1 (bases 1 to 1262)					
NIH-MGC http://mgc.nci.nih.gov/.					
National Institutes of Health, Mammalian Gene Collection (MGC)					
Unpublished (1999)					
Contact: Robert Strausberg, Ph.D.					
Email: cgapbs-remail.nih.gov					
Tissue Procurement: DCTD/DTP					
cDNA Library Preparation: Rubin Laboratory					
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)					
DNA Sequencing by: Agencourt Bioscience Corporation					
Clone distribution: MGC clone distribution information can be					
found through the I.M.A.G.E. Consortium/LLNL at:					
http://image.llnl.gov					
Plate: L1CM2573 row: f column: 09					
High quality sequence start: 46					
High quality sequence stop: 157.					
location/Qualifiers					
1. .1262					
/organism="Homo sapiens"					
/mol_type="mRNA"					
/db_xref="taxon:9606"					
/clone="IMAGE:6382088"					
/tissue_type="carcinoma, cell line"					
/lab_host="DH10B (phage-resistant)"					
/clone_lib="NIH_MGC_40"					
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;					
Site 2: EcoRI; cDNA made by oligo-dT priming.					
Directionally cloned into EcoRI/XhoI sites using the					
following 5' adaptor: GGACGAG(G). Library constructed by					
ling Hong in the laboratory of Gerald M. Rubin (University					

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 16:00:11 : Search time 170.5 Seconds
(without alignments)
1847.304 Million cell updates/sec

Title: US-10-057-136-5
Perfect score: 60
Sequence: 1 ggaaccacgcgcgcgcctgc.....ccgacacgcgcgcgcctccc 60

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*

1: geneseqn19808:*

2: geneseqn19908:*

3: geneseqn20008:*

4: geneseqn2001as:*

5: geneseqn2001bs:*

6: geneseqn2002as:*

7: geneseqn2002bs:*

8: geneseqn2003as:*

9: geneseqn2003bs:*

10: geneseqn2003cs:*

11: geneseqn2003ds:*

12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	60	100.0	60	2	AAV48318	Aav48318 Nucleotid
2	38.2	63.7	60	2	AAV48324	Aav48324 Nucleotid
3	36.6	61.0	156	10	ADK68635	Adk68635 HSP65-MUC
4	36.6	61.0	1818	12	ADF32633	Adf32633 Plasmid J
5	36.6	61.0	2135	12	ADF32629	Adf32629 Plasmid J
6	35.8	59.7	2255	12	AD157667	Ad157667 Human bre
7	35.8	59.7	4144	12	AD157666	Ad157666 Human bre
8	35.2	58.7	1320	2	AAT29190	Aat29190 Glycoprot
9	35	58.3	60	2	AAV48316	Aav48316 Nucleotid
10	35	58.3	180	2	AAT29189	Aat29189 Glycoprot
11	35	58.3	309	1	AAN90579	Aan90579 PDF9.3 CD
12	35	58.3	328	12	AD157670	Ad157670 Human bre
13	35	58.3	1194	12	AD157712	Ad157712 Human bre
14	35	58.3	1378	12	AD157693	Ad157693 Human bre
15	35	58.3	1424	12	ADO23180	Ado23180 Antisense
16	35	58.3	1428	6	ABL60159	Ab160159 Human MUC
17	35	58.3	1428	12	ADO23125	Ado23125 Human MUC
18	35	58.3	1457	12	ADF32627	Adf32627 Plasmid J
19	35	58.3	1527	2	AAV48329	Aav48329 MiniMUC1
20	35	58.3	1614	12	ADK70370	Adk70370 Respirato
21	35	58.3	1630	12	AD157708	Ad157708 Human bre

22	35	58.3	1634	12	AD157689	Ad157689 Human bre
23	35	58.3	1712	12	AD157686	Ad157686 Human bre
24	35	58.3	1721	6	AB876475	Ab876475 CDNA enco
25	35	58.3	1721	6	ABL68032	Ab168032 Ovary can
26	35	58.3	1721	6	ABK09743	Abk09743 Human ova
27	35	58.3	1721	6	ABV94067	Abv94067 Breast ca
28	35	58.3	1721	8	ADA50566	Ada50566 Mucin 1 (
29	35	58.3	1721	8	ACF12906	Acf12906 Human cer
30	35	58.3	1721	9	AAD56838	Aad56838 Human muc
31	35	58.3	1721	12	ADK65926	Adk65926 Human imm
32	35	58.3	1738	12	AD157669	Ad157669 Human bre
33	35	58.3	1755	12	AD157673	Ad157673 Human bre
34	35	58.3	1774	12	ADE43991	Ade43991 Plasmid J
35	35	58.3	1774	12	ADF32625	Adf32625 Plasmid J
36	35	58.3	1799	12	ADO23124	Ado23124 Human MUC
37	35	58.3	1800	10	ADK68629	Adk68629 HSP65-MUC
38	35	58.3	1803	12	AD157699	Ad157699 Human bre
39	35	58.3	1804	6	ABL67539	Ab167539 Thyroid c
40	35	58.3	1804	9	AAD56950	Aad56950 Human muc
41	35	58.3	1804	10	ADD14719	Add14719 Human src
42	35	58.3	1804	12	ADP13294	Adp13294 Renal cel
43	35	58.3	1805	12	ADO28642	Ado28642 Human MUC
44	35	58.3	1808	12	AD157706	Ad157706 Human bre
45	35	58.3	1823	6	AB235228	Ab235228 Human gen

ALIGNMENTS

RESULT 1

AAV48318

ID AAV48318 standard; DNA; 60 BP.

XX

AC AAV48318;

XX

DT 20-NOV-1998 (first entry)

XX

DE Nucleotide sequence encoding MUC1 tandem repeat unit R3.

XX

KW SS; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;

XX

KM tumour; tumour-associated antigen.

XX

OS Homo sapiens.

XX

PN W09837095-A2.

XX

PD 27-AUG-1998.

XX

PF 24-FEB-1998; 98WO-US003693.

XX

PR 24-FEB-1997; 97US-0038253P.

XX

PA (THER-) THERION BIOLOGICS CORP.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (DAND) DANA FARBER CANCER INST INC.

XX

PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;

XX

DR WPI; 1998-467492/40.

XX

PT New recombinant pox virus for tumour therapy - comprises DNA encoding an

PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.

XX

PS Disclosure; Page 11; 42pp; English.

XX

CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an

CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus

CC (RPV). The RPV was used in a pharmaceutical composition also containing

CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The

CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that

CC does not undergo significant genetic deletion, thereby providing an

CC unexpectedly stable and immunogenic pox virus. They can be used to

CC prevent or treat tumours expressing MUC1 tumour-associated antigens

XX Sequence 60 BP; 7 A; 29 C; 19 G; 5 T; 0 U; 0 Other;
SQ

Query Match 100.0%; Score 60; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCTGCGACGAGTGACGTGGCGCCCGACACAGCGCCCGCTCC 60
1 GGATCCACCGCGCGCTGCGACGAGTGACGTGGCGCCCGACACAGCGCCCGCTCC 60
Db

RESULT 2

ID AAV48324 standard; DNA; 60 BP.

XX AAV48324;

DT 20-NOV-1998 (first entry)

DE Nucleotide sequence encoding MUC1 tandem repeat unit R9.

KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
tumour; tumour-associated antigen.

OS Homo sapiens.

PN WO9837095-A2.

PD 27-AUG-1998.

PF 24-FEB-1998; 98WO-US003693.

PR 24-FEB-1997; 97US-0038253P.

PA (THER-) THERION BIOLOGICS CORP.
(USSH) US DEPT HEALTH & HUMAN SERVICES.
(DAND) DANA FARBER CANCER INST INC.

PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;

DR WPI; 1998-467492/40.

PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.

PS Disclosure; Page 11; 42pp; English.

CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
(RPV). The RPV was used in a pharmaceutical composition also containing
an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
does not undergo significant genetic deletion, thereby providing an
unexpectedly stable and immunogenic pox virus. They can be used to
prevent or treat tumours expressing MUC1 tumour-associated antigens

SQ Sequence 60 BP; 12 A; 26 C; 14 G; 8 T; 0 U; 0 Other;

Query Match 63.7%; Score 38.2; DB 2; Length 60;
Best Local Similarity 78.0%; Pred. No. 0.12;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCTGCGACGAGTGACGTGGCGCCCGACACAGCGCCCGCTCC 59
1 GGATCCACCGCGCGCTGCGACGAGTGACGTGGCGCCCGACACAGCGCCCGCTCC 59
Db

RESULT 3

ID ADK68635 standard; cDNA; 156 BP.

XX ADK68635;

XX 06-MAY-2004 (first entry)
DT

DE HSP65-MUC1 antigen CTL epitope related cDNA #3.

KW Human breast cancer; Mycobacterium bovis; heat shock protein 65; BCG;
HSP65; cytotoxic T lymphocyte; CTL; MUC1; HSP65-MUC1 antigen CTL epitope;
ss.

OS Unidentified.

PN CN1368384-A.

PD 11-SEP-2002.

PF 08-FEB-2001; 2001CN-00102614.

PR 08-FEB-2001; 2001CN-00102614.

PA (DIWE-) DIWEIHUAYU BIO TECHNOLOGY CO LTD BEIJING.

PI Yu Y, Li H;

DR WPI; 2003-854662/80.

PT Genetically engineered vaccine of MUC-1 antigen for human breast cancer.

PS Example 2; Page 4 (Disclosure); 14pp; Chinese.

CC The invention relates to a method of preparation of a genetically
engineered vaccine for preventing and treating human breast cancer. The
method comprises fusing the coding gene of the Mycobacterium bovis heat
shock protein 65 (BCG HSP65) with the cytotoxic T lymphocyte (CTL)
epitope gene of MUC1 (antigen cell expressed by human breast cancer
cells) and the HSP65-MUC1 antigen CTL epitope is expressed in Escherichia
coli cells. This sequence represents DNA used in the method of the
invention.

SQ Sequence 156 BP; 37 A; 42 C; 58 G; 19 T; 0 U; 0 Other;

Query Match 61.0%; Score 36.6; DB 10; Length 156;
Best Local Similarity 76.3%; Pred. No. 0.31;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCTGCGACGAGTGACGTGGCGCCCGACACAGCGCCCGCTCC 59
132 GGTTCTACCGCTCCGCGGCTCACGGTGTACCTCTGCTCCGACACCCGTCGCTCC 74
Db

RESULT 4

ID ADF32633 standard; DNA; 1818 BP.

AC ADF32633;

DT 26-FEB-2004 (first entry)

DE Plasmid JNW656 7x VNTR MUC-1 nucleotide sequence.

KW MUC-1 antigen; immune response; MUC-1; variable number of tandem repeat;
VNTR; repeat unit; tumour; metastasis; cytostatic; vaccine; gene therapy;
gene; ds.

OS Synthetic.

PN WO2003100060-A2.

PD 04-DEC-2003.

PF 23-MAY-2003; 2003WO-BP005594.

PR 24-MAY-2002; 2002GB-00012046.

```
PA (GLAX ) GLAXO GROUP LTD.
XX
XX PI Burden N, Ellis JH, Hamblin PA;
XX
XX DR WPI; 2004-042811/04.
XX
PT New nucleic acid molecule encoding a MUC-1 antigen, useful for preparing
PT a composition for treating or preventing tumors or metastases.
XX
XX PS Example; Fig 6; 66pp; English.
XX
CC The present invention describes a nucleic acid molecule which encodes a
CC MUC-1 antigen. The nucleic acid is capable of raising an immune response
CC in vivo, has reduced susceptibility to recombination than full-length MUC
CC -1 and comprises between 1 and 15 variable number of tandem repeats
CC (VNTR) perfect repeat units. Also described: (1) a plasmid comprising the
CC DNA molecule; (2) a protein encoded by the nucleic acid; (3) a
CC pharmaceutical composition comprising the nucleic acid, plasmid or
CC protein and an excipient, diluent or carrier; and (4) a method of
CC treating or preventing tumors or metastases. A MUC1 antigen has
CC cytostatic activity, and can be used in vaccines, and in gene therapy.
CC The nucleic acid is useful for preparing a composition for treating or
CC preventing tumors or metastases. The present sequence is used in the
CC exemplification of the present invention.
XX
SQ Sequence 1818 BP; 346 A; 703 C; 430 G; 339 T; 0 U; 0 Other;

Query Match 61.0%; Score 36.6; DB 12; Length 1818;
Best Local Similarity 76.3%; Pred. No. 0.25;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCCGCTGCGCAGGAGTGACGTGCGGCCCGACACGCGCCCGCTCC 59
Db 583 GGCTCCACCGCCCCCGCCAGCCACGGGTGTCACTCGGCCCGGACACCAAGCCGCCCC 641

RESULT 5
ADF32629
ID ADF32629 standard; DNA; 2135. BP.
XX
XX AC ADF32629;
XX
XX DT 26-FEB-2004 (first entry)
XX
XX DE Plasmid JNW319 7x VNTR MUC-1 nucleotide sequence.
XX
XX KM MUC-1 antigen; immune response; MUC-1; variable number of tandem repeat;
XX VNTR; repeat unit; tumour; metastasis; cytostatic; vaccine; gene therapy;
XX gene; ds.
XX
XX OS Synthetic.
XX
XX PN WO2003100060-A2.
XX
XX PD 04-DEC-2003.
XX
XX PF 23-MAY-2003; 2003WO-EP005594.
XX
XX PR 24-MAY-2002; 2002GB-00012046.
XX
XX PA (GLAX ) GLAXO GROUP LTD.
XX
XX PI Burden N, Ellis JH, Hamblin PA;
XX
XX DR WPI; 2004-042811/04.
XX
XX PT New nucleic acid molecule encoding a MUC-1 antigen, useful for preparing
XX a composition for treating or preventing tumors or metastases.
XX
XX PS Example; Fig 4; 66pp; English.
XX
XX CC The present invention describes a nucleic acid molecule which encodes a
XX MUC-1 antigen. The nucleic acid is capable of raising an immune response
```

```
CC in vivo, has reduced susceptibility to recombination than full-length MUC
CC -1 and comprises between 1 and 15 variable number of tandem repeats
CC (VNTR) perfect repeat units. Also described: (1) a plasmid comprising the
CC DNA molecule; (2) a protein encoded by the nucleic acid; (3) a
CC pharmaceutical composition comprising the nucleic acid, plasmid or
CC protein and an excipient, diluent or carrier; and (4) a method of
CC treating or preventing tumors or metastases. A MUC1 antigen has
CC cytostatic activity, and can be used in vaccines, and in gene therapy.
CC The nucleic acid is useful for preparing a composition for treating or
CC preventing tumors or metastases. The present sequence is used in the
CC exemplification of the present invention.
XX
XX SQ Sequence 2135 BP; 395 A; 812 C; 524 G; 404 T; 0 U; 0 Other;

Query Match 61.0%; Score 36.6; DB 12; Length 2135;
Best Local Similarity 76.3%; Pred. No. 0.24;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCCGCTGCGCAGGAGTGACGTGCGGCCCGACACGCGCCCGCTCC 59
Db 642 GGCTCCACCGCCCCCGCCAGCCACGGGTGTCACTCGGCCCGGACACCAAGCCGCCCC 700

RESULT 6
AD157667
ID AD157667 standard; cDNA; 2255 BP.
XX
XX AC AD157667;
XX
XX DT 22-APR-2004 (first entry)
XX
XX DE Human breast specific nucleic acid (BSNA) #38.
XX
XX KM Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
XX breast cancer; cytostatic.
XX
XX OS Homo sapiens.
XX
XX PN WO2003106648-A2.
XX
XX PD 24-DEC-2003.
XX
XX PF 16-JUN-2003; 2003WO-US018934.
XX
XX PR 14-JUN-2002; 2002US-0389327P.
XX
XX PA (DIAD-) DIADEXUS INC.
XX
XX PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
XX
XX DR WPI; 2004-082185/08.
XX
XX DR P-PSDB; AD157741.
XX
XX PT Novel isolated polypeptide comprising breast specific protein sequences,
XX useful for diagnosing or monitoring presence and metastases of breast
XX cancer in patient.
XX
XX PS Claim 1; SEQ ID NO 38; 370pp; English.
XX
XX CC The invention relates to human breast specific nucleic acids (BSNA) and
XX the breast specific proteins (BSP) they encode. The nucleic acids are
XX useful for determining the presence of a BSNA in a sample which involves
XX contacting the sample with a BSNA under conditions in which the BSNA will
XX selectively hybridise to a BSNA in the sample, and detecting the
XX hybridisation. The nucleic acids are useful for determining the presence
XX of a BSP in a sample which involves contacting the sample with suitable
XX reagent under conditions in which the reagent will selectively interact
XX with the BSP, and detecting the interaction of the reagent with a BSP in
XX the sample. The nucleic acids and proteins are useful for diagnosing or
XX monitoring the presence and metastases of breast cancer in a patient,
XX which involves determining an amount of nucleic acid or protein and
XX comparing the determined amount of nucleic acid or protein in the sample
XX of the patient to the amount of a breast specific marker in a normal
```

CC	control, where a difference in the determined amount in the sample
CC	compared to the amount in the control is associated with the presence of
CC	breast cancer. The sequences are useful for treating a patient with
CC	breast cancer, involving administering a composition consisting of a BSNA
CC	or a BSP to a patient, where the administration induces an immune
CC	response against the breast cancer cell expressing the BSNA or BSP. This
CC	sequence represents a human BSNA of the invention.
XX	
SQ	Sequence 2255 BP; 442 A; 763 C; 585 G; 461 T; 0 U; 4 Other;
	Query Match 59.7%; Score 35.8; DB 12; Length 2255;
	Best Local Similarity 72.9%; Pred. No. 0.4;
	Matches 43; Conservative 2; Mismatches 14; Indels 0; Gaps
OY	
Db	1 GGATCACCGCGCCGCTGCACAGGAGTGACGTGGCGCCCGACACGCCCGCGCTTC 59 : 578 GGCTCACCGCSCCCSCAGCCCAAGGTGTCTACTCTGGCCCCCGACACGCGCGCCC 636
RESULT 7	
ID	ADIS7666 standard; cDNA; 4144 BP.
XX	
AC	ADIS7666;
XX	
DT	22-APR-2004 (first entry)
XX	
DE	Human breast specific nucleic acid (BSNA) #37.
XX	
KM	Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
KW	breast cancer; cytosstatic.
XX	
OS	Homo sapiens.
XX	
PN	WO2003106648-A2.
XX	
PD	24-DEC-2003.
XX	
PF	16-JUN-2003; 2003WO-US018934.
XX	
PR	14-JUN-2002; 2002US-0389327P.
XX	
PA	(DIAD-) DIADEXUS INC.
XX	
PI	Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
XX	
DR	WPI; 2004-082185/08.
DR	P-PSDB; ADIS7740.
XX	
PT	Novel isolated polypeptide comprising breast specific protein sequences,
PT	useful for diagnosing or monitoring presence and metastases of breast
PT	cancer in patient.
XX	
PS	Claim 1; SEQ ID NO 37; 370bp; English.
XX	
CC	The invention relates to human breast specific nucleic acids (BSNA) and
CC	the breast specific proteins (BSP) they encode. The nucleic acids are
CC	useful for determining the presence of a BSNA in a sample which involves
CC	contacting the sample with a BSNA under conditions in which the BSNA will
CC	selectively hybridise to a BSNA in the sample, and detecting the
CC	hybridisation. The nucleic acids are useful for determining the presence
CC	of a BSP in a sample which involves contacting the sample with suitable
CC	reagent under conditions in which the reagent will selectively interact
CC	with the BSP, and detecting the interaction of the reagent with a BSP in
CC	the sample. The nucleic acids and proteins are useful for diagnosing or
CC	monitoring the presence and metastases of breast cancer in a patient,
CC	which involves determining an amount of nucleic acid or protein in the sample
CC	comparing the determined amount of nucleic acid or protein in the sample
CC	of the patient to the amount of a breast specific marker in a normal
CC	control, where a difference in the determined amount in the sample
CC	compared to the amount in the control is associated with the presence of
CC	breast cancer. The sequences are useful for treating a patient with
CC	breast cancer, involving administering a composition consisting of a BSNA

CC	or a BSP to a patient, where the administration induces an immune
CC	response against the breast cancer cell expressing the BSNA or BSP. This
CC	sequence represents a human BSNA of the invention.
XX	
SQ	Sequence 4144 BP; 634 A; 1896 C; 1058 G; 552 T; 0 U; 4 Other;
Query Match	59.7%; Score 35.8; DB 12; Length 4144;
Best Local Similarity	72.9%; Pred. No. 0.38;
Matches 43; Conservative	2; Mismatches 14; Indels 0; Gaps 0
OY	1 GGATCCACCGCGGCGCTGTGGCAGCAGATGACGTCGGCGGCCGCACACCGCGCCGCTCC 59
Db	633 GGCTCCACCGCGSCCCSCAGCCCCACGATGACACTCGGCCCCGACACCAGCGCGCCC 691
RESULT 8	
AAT29190	standard; cDNA to mRNA; 1320 BP.
XX	
AC	AAT29190;
XX	
DT	26-JUL-1996 (first entry)
XX	
DE	Glycoprotein 39 3' fragment.
XX	
KW	Glycoprotein 39; gp39; lambda gt11 cDNA library; gastric cancer;
KW	cell line KATO-III; tumour; immune abnormality; marker;
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..984
FT	/tag= a
FT	polyA_signal
FT	1267..1272
FT	/tag= b
FT	polyA_signal
FT	1293..1320
FT	/tag= c
XX	
PN	JP07051065-A.
XX	
PD	28-FEB-1995.
XX	
PF	21-FEB-1992; 92JP-00035085.
XX	
PR	21-FEB-1992; 92JP-00035085.
XX	
PA	(NIKO-) NIPPON KOTAI KENKYUSHO KK.
PA	(UYKA-) UNIV KAGOSHIMA.
DR	WPI; 1995-127356/17.
DR	P-PADB; AAR96298.
XX	
PT	Glyco:protein 39 gene - used in the mass production of glyco:protein 39,
PT	for use as tumour and immune abnormality marker.
XX	
PS	Claim 2; Page 9-10; 10pp; Japanese.
XX	
CC	The sequences given in AAT29189-90 encode portions of glycoprotein 39.
CC	This sequence represents the 3' portion of the glycoprotein 39 gene and
CC	was isolated from lambda gt11 cDNA library derived from the gastric
CC	cancer cell line KATO-III. Glycoprotein 39 is expected to be used as a
CC	tumour marker, an immune abnormality marker or a marker for various
CC	inflammatory diseases
XX	
SQ	Sequence 1320 BP; 288 A; 417 C; 317 G; 298 T; 0 U; 0 Other;
Query Match	58.7%; Score 35.2; DB 2; Length 1320;
Best Local Similarity	76.8%; Pred. No. 0.62;
Matches 43; Conservative	0; Mismatches 13; Indels 0; Gaps 0;
OY	1 GGATCCACCGCGCGCTGTGGCAGCAGATGACGTCGGCGGCCGCACACCGCGCCCGC 56


```
Db      1 GGCCTCCACCGCCCCCGCCAGCCAGGTGTCACTCGGCCCCGGACACACGAGCCCGC 56
RESULT 9
AAV48316
ID      AAV48316 standard; cDNA; 60 BP.
XX
AC      AAV48316;
XX
DT      20-NOV-1998 (first entry)
XX
DE      Nucleotide sequence encoding MUC1 tandem repeat unit.
XX
KM      ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
XX      tumour; tumour-associated antigen.
XX
OS      Homo sapiens.
XX
FH      Key
FT      CDS
FT      1.60
FT      /*tag= a
FT      /product= "MUC1 tandem repeat unit"
XX
XX      WO9837095-A2.
XX
XX      27-AUG-1998.
XX
XX      24-FEB-1998; 98WO-US003693.
XX
XX      24-FEB-1997; 97US-0038253P.
XX
XX      (THER-) THERION BIOLOGICS CORP.
XX      (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX      (DAND ) DANA FABER CANCER INST INC.
XX
XX      Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX
XX      WPI; 1998-467492/40.
XX      P-PSDB; AAW77229.
XX
XX      New recombinant pox virus for tumour therapy - comprises DNA encoding an
XX      immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
XX      Example 1; Page 20; 42pp; English.
XX
XX      The MUC1 tandem repeat unit was used to create an immunogenic mini-MUC1
XX      fragment for inclusion in a recombinant pox virus (RPV). The RPV was used
XX      in a pharmaceutical composition also containing an immunomodulator to
XX      generate MUC1 specific cytotoxic T-lymphocytes. The recombinant pox virus
XX      therefore encodes an immunogenic MUC1 fragment that does not undergo
XX      significant genetic deletion, thereby providing an unexpectedly stable
XX      and immunogenic pox virus. They can be used to prevent or treat tumours
XX      expressing MUC1 tumour-associated antigens
XX
XX      Sequence 60 BP; 7 A; 33 C; 16 G; 4 T; 0 U; 0 Other;
SQ
Query Match 58.3%; Score 35; DB 2; Length 60;
Best Local Similarity 74.6%; Pred. No. 0.93;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY      1 GGATCCACCGCGCGCCTTGGCAGAGTGACGTGCGCGCCGACACGCGCCCGCTCC 59
Db      1 GGCCTCCACCGCCCCCGCCAGCCAGGTGTCACTCGGCCCCGGACACACGAGCCCGC 59
RESULT 10
AAT29189
ID      AAT29189 standard; cDNA to mRNA; 180 BP.
XX
AC      AAT29189;
XX
DT      25-JUL-1996 (first entry)
XX
```

```
DE      Glycoprotein 39 5' fragment.
XX
XX      Glycoprotein 39; gp39; lambda gt11 cDNA library; gastric cancer;
XX      cell line KATO-III; tumour; immune abnormality; marker;
XX      inflammatory disease; ds.
XX
XX      Homo sapiens.
XX
XX      Key
XX      repeat_region 1.180
XX      repeat_unit 1.60
XX      /*tag= a
XX      /*tag= b
XX
XX      JP07051065-A.
XX
XX      28-FEB-1995.
XX
XX      21-FEB-1992; 92JP-00035085.
XX
XX      21-FEB-1992; 92JP-00035085.
XX
XX      (NIKO-) NIPPON KOTAI KENKYUSHO KK.
XX      (UYKA-) UNIV KAGOSHIMA.
XX
XX      WPI; 1995-127356/17.
XX      P-PSDB; AAR96297.
XX
XX      Glyco:protein 39 gene - used in the mass production of glyco:protein 39,
XX      for use as tumour and immune abnormality marker.
XX
XX      Claim 2; Page 8; 10pp; Japanese.
XX
XX      The sequences given in AAT29189-90 encode portions of glycoprotein 39.
XX      This sequence represents the 5' portion of the glycoprotein 39 gene and
XX      was isolated from lambda gt11 cDNA library derived from the gastric
XX      cancer cell line KATO-III. Glycoprotein 39 is expected to be used as a
XX      tumour marker, an immune abnormality marker or a marker for various
XX      inflammatory diseases
XX
XX      Sequence 180 BP; 21 A; 91 C; 56 G; 12 T; 0 U; 0 Other;
SQ
Query Match 58.3%; Score 35; DB 2; Length 180;
Best Local Similarity 74.6%; Pred. No. 0.84;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY      1 GGATCCACCGCGCGCCTTGGCAGAGTGACGTGCGCGCCGACACGCGCCCGCTCC 59
Db      121 GGCCTCCACCGCGCGCCGACGCCAGGTGTCACTCGGCCCCGGACACACGAGCCCGC 179
RESULT 11
AAN90579/c
ID      AAN90579 standard; cDNA; 309 BP.
XX
AC      AAN90579;
XX
XX      27-AUG-2003 (revised)
XX      25-MAR-2003 (revised)
XX      04-DEC-1989 (first entry)
XX
XX      pDF9.3 cDNA insert.
XX
XX      pDF9.3; human DF3 breast carcinoma-associated antigen epitope.
XX
XX      Human MCF-7 breast carcinoma cells; 'ATCC HTB22'.
XX
XX      WO8907107-A.
XX      10-AUG-1989.
XX
XX      29-JAN-1988; 88US-00149831.
XX
XX
```

PR 29-JAN-1988; 88US-00149831.
XX
PA (DANA-) DANA-FARBER CANCER.
XX
PI Kufe DW;
XX
DR WPI; 1989-248989/34.
DR P-PSDB; AAP91045, AAP91053, AAP91054, AAP90146.
XX
PT Recombinant polypeptide(s) - contains DF 3 breast carcinoma antigen
PT epitope and useful as assay reagents, and encoding DNA sequences.
XX
PS Claim 1; Fig 4; 31pp; English.
XX
CC The sequence encodes a carbohydrate-free polypeptide contg. DF3 breast
CC carcinoma antigen epitope. Useful as a competitive binding assay reagent
CC and improves diagnosis. The CDNA consists of nearly identical 60 BP
CC tandem repeats which are 85% GC-rich. See also AAP91045, AAP91046,
CC AAP91053, AAP91054. (Updated on 25-MAR-2003 to correct PA field.)
CC (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 309 BP; 21 A; 92 C; 160 G; 36 T; 0 U; 0 Other;

Query Match 58.3%; Score 35; DB 1; Length 309;
Best Local Similarity 74.6%; Pred. No. 0.8;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCTGCGCAGGAGTGACGTGCGCGCCGACACAGCGCCCGCTCC 59
Db 266 GGCTCCACCGCGCGCCCGCAGCCACGAGTGTCACCTCGGCCCGACACAGCGCGCCCG 208

RESULT 12
ADIS7670
ID ADIS7670 standard; cDNA; 328 BP.
XX
AC ADIS7670;
XX
DT 22-APR-2004 (first entry)
XX
DE Human breast specific nucleic acid (BSNA) #41.
XX
KW Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
KM breast cancer; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2003106648-A2.
XX
PD 24-DEC-2003.
XX
PF 16-JUN-2003; 2003WO-US018934.
XX
PR 14-JUN-2002; 2002US-0389327P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
XX
DR WPI; 2004-082185/08.
DR P-PSDB; ADIS7744.
XX
PT Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.
XX
PS Claim 1; SEQ ID NO 41; 370pp; English.
XX
CC The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the

CC hybridisation. The nucleic acids are useful for determining the presence
CC of a BSP in a sample which involves contacting the sample with suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein and
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.
XX
SQ Sequence 328 BP; 47 A; 124 C; 95 G; 61 T; 0 U; 1 Other;

Query Match 58.3%; Score 35; DB 12; Length 328;
Best Local Similarity 74.6%; Pred. No. 0.8;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCGCTGCGCAGGAGTGACGTGCGCGCCGACACAGCGCCCGCTCC 59
Db 95 GGCTCCACCGCGCGCGCAGCCACGAGTGTCACCTCGGCCCGACACAGCGCGCCCG 153

RESULT 13
ADIS7712
ID ADIS7712 standard; cDNA; 1194 BP.
XX
AC ADIS7712;
XX
DT 22-APR-2004 (first entry)
XX
DE Human breast specific nucleic acid (BSNA) #83.
XX
KW Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
KM breast cancer; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2003106648-A2.
XX
PD 24-DEC-2003.
XX
PF 16-JUN-2003; 2003WO-US018934.
XX
PR 14-JUN-2002; 2002US-0389327P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
XX
DR WPI; 2004-082185/08.
DR P-PSDB; ADIS7782.
XX
PT Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.
XX
PS Claim 1; SEQ ID NO 83; 370pp; English.
XX
CC The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the
CC hybridisation. The nucleic acids are useful for determining the presence
CC of a BSP in a sample which involves contacting the sample with suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in

CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein and
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.

XX Sequence 1194 BP; 243 A; 438 C; 275 G; 238 T; 0 U; 0 Other;

Query Match 58.3%; Score 35; DB 12; Length 1194;
Best Local Similarity 74.6%; Pred. No. 0.71;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY 1 GGATCCACCGCGCGCGCTGGCGACGAGTGACGTGGCGCCCGACACGCGCCCGCTCC 59
|||
DB 596 GGCTCCACCGCGCGCGCGCGCTGGCGACGAGTGACGTGGCGCCCGACACGCGCGCCCG 654

RESULT 14

AD157693
ID AD157693 standard; cDNA; 1378 BP.

XX AC AD157693;

XX DT 22-APR-2004 (first entry)

XX DE Human breast specific nucleic acid (BSNA) #64.

XX DE Human breast specific nucleic acid; BSNA; gene; ss; metastasis;
XX KW breast cancer; cytostatic.

XX OS Homo sapiens.

XX PN WO2003106648-A2.

XX PD 24-DEC-2003.

XX PF 16-JUN-2003; 2003WO-US018934.

XX PR 14-JUN-2002; 2002US-0389327P.

XX PA (DIAD-) DIADEXUS INC.

XX PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;

XX DR WPI; 2004-082185/08.

XX DR P-PSDB; AD157765.

XX PT Novel isolated polypeptide comprising breast specific protein sequences,
XX PT useful for diagnosing or monitoring presence and metastases of breast
XX PT cancer in patient.

XX PS Claim 1; SEQ ID NO 64; 370pp; English.

XX CC The invention relates to human breast specific nucleic acids (BSNA) and
XX CC the breast specific proteins (BSP) they encode. The nucleic acids are
XX CC useful for determining the presence of a BSNA in a sample which involves
XX CC contacting the sample with a BSNA under conditions in which the BSNA will
XX CC selectively hybridise to a BSNA in the sample, and detecting the
XX CC hybridisation. The nucleic acids are useful for determining the presence
XX CC of a BSP in a sample which involves contacting the sample with suitable
XX CC reagent under conditions in which the reagent will selectively interact
XX CC with the BSP, and detecting the interaction of the reagent with a BSP in
XX CC the sample. The nucleic acids and proteins are useful for diagnosing or
XX CC monitoring the presence and metastases of breast cancer in a patient,
XX CC which involves determining an amount of nucleic acid or protein and
XX CC comparing the determined amount of nucleic acid or protein in the sample

CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.

XX Sequence 1378 BP; 295 A; 489 C; 316 G; 278 T; 0 U; 0 Other;

Query Match 58.3%; Score 35; DB 12; Length 1378;
Best Local Similarity 74.6%; Pred. No. 0.7;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY 1 GGATCCACCGCGCGCGCTGGCGACGAGTGACGTGGCGCCCGACACGCGCCCGCTCC 59
|||
DB 596 GGCTCCACCGCGCGCGCGCGCTGGCGACGAGTGACGTGGCGCCCGACACGCGCGCCCG 654

RESULT 15
ADO23180/C
ID ADO23180 standard; RNA; 1424 BP.

XX AC ADO23180;

XX DT 12-AUG-2004 (first entry)

XX DE Antisense human MUC1 mucin glycoprotein RNA (coding sequence) SeqID 75.

XX DE human; MUC1; mucin glycoprotein; cancer; chemotherapeutic; MUC1/BCD; ss;
XX KW antisense.

XX OS Homo sapiens.

XX PN WO200404160-A2.

XX PD 27-MAY-2004.

XX PF 12-NOV-2003; 2003WO-US035848.

XX PR 13-NOV-2002; 2002US-00293391.

XX PR 29-MAY-2003; 2003US-00447839.

XX PA (DAND) DANA FARBER CANCER INST INC.

XX PA (ILEX-) ILEX PROD INC.

XX PI Kufe DW, Kharbanda S, Weltman SD;

XX DR WPI; 2004-420304/39.

XX PT Double-stranded RNA complex useful for inhibiting proliferation of cancer
XX PT cell expressing MUC1 mucin glycoprotein, comprises first and second RNA
XX PT sequences.

XX PS Disclosure; SEQ ID NO 75; 112pp; English.

XX CC This invention relates to novel modulators of the human MUC1 mucin
XX CC glycoprotein for use in cancer therapeutics, where MUC1 is a protein that
XX CC acts to inhibit the apoptotic response to genotoxic stress caused by
XX CC chemotherapeutic agents. In particular, it refers to modulators of the
XX CC MUC1 extracellular domain (MUC1/BCD). The method refers to using double-
XX CC stranded RNA complexes as MUC1 interference RNA compositions such that
XX CC MUC1 expression is inhibited, which in turn inhibits cancer cell
XX CC proliferation. The present invention describes screening assays to
XX CC identify compounds that inhibit the binding of various MUC1 ligands such
XX CC as neuregulin 2, as well as agonists, antagonists and antibodies thereof.
XX CC Furthermore, it provides MUC1 antisense, small interfering RNA (siRNA)
XX CC and small molecules in combination with chemotherapeutic agents that are
XX CC useful in the field of cancer therapy. This polynucleotide sequence is
XX CC the antisense human MUC1 RNA of the invention.

XX Sequence 1424 BP; 309 A; 325 C; 495 G; 0 T; 295 U; 0 Other;

Query Match 58.3%; Score 35; DB 12; Length 1424;
Best Local Similarity 74.6%; Pred. No. 0.7;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCCCTGCGCACGGAGTGACGTGCGCGCGCGACACAGCGCGCGCTCC 59
Db 1041 GGCTCCACCGCGCGCGCCCGACAGCCCGGTGTCACTCGGCGCGGACACACAGCGCGCGCC 983

Search completed: January 15, 2005, 20:36:07
Job time : 170.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 22:37:00 ; Search time 183.4 Seconds
(without alignments)
1879.789 Million cell updates/sec

Title: US-10-057-136-4

Perfect score: 60
Sequence: 1 ggcagctactgcaccaccgcgc.....ctgatacaagaccctgcacct 60

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptoddata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptoddata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptoddata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptoddata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptoddata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptoddata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptoddata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptoddata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptoddata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptoddata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptoddata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptoddata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptoddata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptoddata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptoddata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptoddata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptoddata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptoddata/1/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptoddata/1/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptoddata/1/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptoddata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	100.0	60	US-10-057-136-4	Sequence 4, Appli
2	60	100.0	2297	US-10-406-317-41	Sequence 41, Appl
3	37.6	62.7	60	US-10-057-136-12	Sequence 12, Appl
4	35.4	59.0	60	US-10-057-136-10	Sequence 10, Appl
5	33.4	55.7	60	US-10-057-136-8	Sequence 8, Appli
6	33.4	55.7	60	US-10-057-136-9	Sequence 9, Appli
7	33.2	55.3	60	US-10-057-136-7	Sequence 7, Appli
8	29	48.3	60	US-10-057-136-11	Sequence 11, Appl
9	28.6	47.7	1230	US-10-369-493-42059	Sequence 42059, A
10	28	46.7	780	US-10-437-963-15324	Sequence 15324, A
11	27.4	45.7	60	US-10-057-136-6	Sequence 6, Appli
12	27.4	45.7	1682	US-10-479-638-10	Sequence 10, Appl

13	27	45.0	60	14	US-10-057-136-2	Sequence 2, Appli
14	27	45.0	78	14	US-10-057-136-13	Sequence 13, Appl
15	27	45.0	1424	16	US-10-447-839A-75	Sequence 75, Appl
16	27	45.0	1428	16	US-10-447-839A-20	Sequence 20, Appl
17	27	45.0	1527	14	US-10-057-136-19	Sequence 19, Appl
18	27	45.0	1799	16	US-10-447-839A-19	Sequence 19, Appl
19	27	45.0	1804	9	US-09-964-824A-573	Sequence 573, App
20	27	45.0	1804	15	US-10-029-517-17	Sequence 17, Appl
21	27	45.0	1804	17	US-10-717-597-30	Sequence 30, Appl
22	27	45.0	1804	17	US-10-775-920-84	Sequence 84, Appl
23	27	45.0	1823	15	US-10-101-510-339	Sequence 339, App
24	27	45.0	4139	9	US-09-964-824A-105	Sequence 105, App
25	27	45.0	4139	9	US-09-964-824A-578	Sequence 578, App
26	27	45.0	4139	9	US-09-864-864-334	Sequence 334, App
27	27	45.0	4139	9	US-09-880-107-2121	Sequence 2121, Ap
28	27	45.0	4139	11	US-09-968-007A-751	Sequence 751, App
29	27	45.0	4139	14	US-10-171-311-157	Sequence 157, App
30	27	45.0	4139	15	US-10-177-293-310	Sequence 310, App
31	27	45.0	4139	16	US-10-440-464-155	Sequence 155, App
32	27	45.0	4139	17	US-10-734-564-53	Sequence 53, Appl
33	27	45.0	4139	17	US-10-775-920-80	Sequence 80, Appl
34	27	45.0	4139	17	US-10-775-920-85	Sequence 85, Appl
35	27	45.0	8181	16	US-10-447-839A-18	Sequence 18, Appl
36	27	45.0	8186	14	US-10-247-703-23	Sequence 23, Appl
37	27	45.0	8186	15	US-10-029-517-19	Sequence 19, Appl
38	26.4	44.0	389	10	US-09-918-995-37681	Sequence 37681, A
39	26.4	44.0	922	16	US-10-424-599-61938	Sequence 61938, A
40	26.4	44.0	3722	16	US-10-305-720-1129	Sequence 1129, Ap
41	26.2	43.7	596	18	US-10-425-115-28374	Sequence 28374, A
42	26	43.3	1194	15	US-10-156-761-5206	Sequence 5206, Ap
43	26	43.3	9025608	15	US-10-156-761-1	Sequence 1, Appli
44	25.6	42.7	60	14	US-10-057-136-14	Sequence 14, Appl
45	25.6	42.7	518	14	US-10-247-703-38	Sequence 38, Appl

ALIGNMENTS

RESULT 1
US-10-057-136-4
; Sequence 4, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHIOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KOFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: TUMOR-ASSOCIATED ANTIGEN
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-4

Query Match 100.0%; Score 60; DB 14; Length 60;
Best Local Similarity 100.0%; Pred. No. 3e-11;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCAGTACTGCACACCGGCACATGGCGTAACATCAGCACCTGATACAGAGACTGCACCT 60
|||||

Db 1 GGCAGTACTGCACCGGCACATGGCGTAACATCAGACCTGTATACAAGACCTGCACCT 60

RESULT 2

US-10-406-317-41
; Sequence 41, Application US/10406317
; Publication No. US20040019195A1
; GENERAL INFORMATION:
; APPLICANT: Schlom, Jeffrey;
; APPLICANT: Hodge, James;
; APPLICANT: Panicali, Dennis
; TITLE OF INVENTION: A recombinant vector expressing multiple constimulatory
; FILE REFERENCE: 38163-0189
; CURRENT APPLICATION NUMBER: US/10/406,317
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US/09/856,988
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: PCT/US99/26866
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/111,582
; PRIOR FILING DATE: 1998-12-09
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: VECTOR
US-10-406-317-41

Query Match 100.0%; Score 60; DB 16; Length 2297;
Best Local Similarity 100.0%; Pred. No. 5.9e-11;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCAGTACTGCACCGGCACATGGCGTAACATCAGACCTGTATACAAGACCTGCACCT 60
Db 406 GGCAGTACTGCACCGGCACATGGCGTAACATCAGACCTGTATACAAGACCTGCACCT 465

RESULT 3

US-10-057-136-12
; Sequence 12, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-12

Query Match 62.7%; Score 37.6; DB 14; Length 60;
Best Local Similarity 76.7%; Pred. No. 0.0018;

Matches 46; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 GGCAGTACTGCACCGGCACATGGCGTAACATCAGACCTGTATACAAGACCTGCACCT 60
Db 1 GGTAGTACAGCGCCACCGGCACATGGCGTACAGCGGCTCCGATACGAGACCGGCGCCT 60

RESULT 4

US-10-057-136-10
; Sequence 10, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: TUMOR-ASSOCIATED ANTIGEN
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-10

Query Match 59.0%; Score 35.4; DB 14; Length 60;
Best Local Similarity 79.2%; Pred. No. 0.01;
Matches 42; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 7 ACTGACACCGGCACATGGCGTAACATCAGACCTGTATACAAGACCTGCACCT 59
Db 7 ACTGCCCTCCGCGCATGTGTGACTCCTCCTGACACAGGCCAGCCCC 59

RESULT 5

US-10-057-136-8
; Sequence 8, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-8

Query Match 55.7%; Score 33.4; DB 14; Length 60;
Best Local Similarity 72.9%; Pred. No. 0.052;
Matches 43; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

OY 1 GGCAGTACTGCACCGCAGCATGGCGTAACATCAGCAGCCTGATACAGACCTGCACC 59
Db 1 GGCAGCAGCGCAGCCGCCGACACAGCGGGTCACAAGCGGCCGACAGACTGCAGCTGCGCC 59

RESULT 6
US-10-057-136-9
Sequence 9, Application US/10057136
Publication No. US20030021770A1
GENERAL INFORMATION:
APPLICANT: SCHLOM, JEFFREY
APPLICANT: KANTOR, JUDITH
APPLICANT: KUFE, DONALD
APPLICANT: PANICALI, DENNIS
APPLICANT: GRITZ, LINDA
TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
TITLE OF INVENTION: TUMOR-ASSOCIATED ANTIGEN
FILE REFERENCE: 700953/47113C
CURRENT APPLICATION NUMBER: US/10/057,136
CURRENT FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: 09/366,670
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: PCT/US98/03693
PRIOR FILING DATE: 1998-02-24
PRIOR APPLICATION NUMBER: 60/038,253
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
LENGTH: 60
TYPE: DNA
ORGANISM: Homo sapiens
US-10-057-136-9

Query Match 55.7%; Score 33.4; DB 14; Length 60;
Best Local Similarity 72.9%; Pred. No. 0.052;
Matches 43; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

OY 1 GGCAGTACTGCACCGCAGCATGGCGTAACATCAGCAGCCTGATACAGACCTGCACC 59
Db 1 GGAAGTACCGCTCCACCTGCACACGCGGTCACAAGCGGCCAGACACTGCAGCTGCGCC 59

RESULT 7
US-10-057-136-7
Sequence 7, Application US/10057136
Publication No. US20030021770A1
GENERAL INFORMATION:
APPLICANT: SCHLOM, JEFFREY
APPLICANT: KANTOR, JUDITH
APPLICANT: KUFE, DONALD
APPLICANT: PANICALI, DENNIS
APPLICANT: GRITZ, LINDA
TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
TITLE OF INVENTION: TUMOR-ASSOCIATED ANTIGEN
FILE REFERENCE: 700953/47113C
CURRENT APPLICATION NUMBER: US/10/057,136
CURRENT FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: 09/366,670
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: PCT/US98/03693
PRIOR FILING DATE: 1998-02-24
PRIOR APPLICATION NUMBER: 60/038,253
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 60

TYPE: DNA
ORGANISM: Homo sapiens
US-10-057-136-7

Query Match 55.3%; Score 33.2; DB 14; Length 60;
Best Local Similarity 75.9%; Pred. No. 0.061;
Matches 41; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

OY 7 ACTGACCAACCGGACATGGCGTAACATCAGCAGCCTGATACAGACCTGCACCT 60
Db 7 ACGGCCCCCTGCTCAGGTGTACATCCGCCCGGATACAGACCTGCGCCCT 60

RESULT 8
US-10-057-136-11
Sequence 11, Application US/10057136
Publication No. US20030021770A1
GENERAL INFORMATION:
APPLICANT: SCHLOM, JEFFREY
APPLICANT: KANTOR, JUDITH
APPLICANT: KUFE, DONALD
APPLICANT: PANICALI, DENNIS
APPLICANT: GRITZ, LINDA
TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
TITLE OF INVENTION: TUMOR-ASSOCIATED ANTIGEN
FILE REFERENCE: 700953/47113C
CURRENT APPLICATION NUMBER: US/10/057,136
CURRENT FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: 09/366,670
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: PCT/US98/03693
PRIOR FILING DATE: 1998-02-24
PRIOR APPLICATION NUMBER: 60/038,253
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 60
TYPE: DNA
ORGANISM: Homo sapiens
US-10-057-136-11

Query Match 48.3%; Score 29; DB 14; Length 60;
Best Local Similarity 71.7%; Pred. No. 1.7;
Matches 38; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY 7 ACTGACCAACCGGACATGGCGTAACATCAGCAGCCTGATACAGACCTGCACC 59
Db 7 ACGGACCTCCAGCAGCAGAGTCACTGTGACCCGACACCCCTCCAGCTCC 59

RESULT 9
US-10-369-493-42059/C
Sequence 42059, Application US/10369493
Publication No. US2003023675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 42059
LENGTH: 1230
TYPE: DNA
ORGANISM: Lactococcus lactis

Query Match 45.0%; Score 27; DB 14; Length 60;
Best Local Similarity 66.1%; Pred. No. 8.6;
Matches 39; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 GGCACTGACCGACCGGACATGGCGTAACATCAGACCTGATACAGACCTGCACC 59
DB 1 GGCTCCACCGCCCCCGACCGGATGTCACCTCGGCCCGGACACGAGCGGCCCC 59

RESULT 14

US-10-057-136-13
; Sequence 13, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; TITLE OF INVENTION: TUMOR-ASSOCIATED ANTIGEN
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 78
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-13

Query Match 45.0%; Score 27; DB 14; Length 78;
Best Local Similarity 66.1%; Pred. No. 9;
Matches 39; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 GGCACTGACCGACCGGACATGGCGTAACATCAGACCTGATACAGACCTGCACC 59
DB 1 GGCTCCACCGCCCCCGACCGGATGTCACCTCGGCCCGGACACGAGCGGCCCC 59

RESULT 15

US-10-447-839A-75/c
; Sequence 75, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald W.
; APPLICANT: Kharbanda, Surender
; APPLICANT: Weltman, Steven D.
; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT APPLICATION NUMBER: US/10/447,839A
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 75
; LENGTH: 1424
; TYPE: RNA
; ORGANISM: ARTIFICIAL
; FEATURE:

; OTHER INFORMATION: Synthesized Sequence
US-10-447-839A-75

Query Match 45.0%; Score 27; DB 16; Length 1424;
Best Local Similarity 66.1%; Pred. No. 16;
Matches 39; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 GGCACTGACCGACCGGACATGGCGTAACATCAGACCTGATACAGACCTGCACC 59
DB 1041 GGCTCCACCGCCCCCGACCGGATGTCACCTCGGCCCGGACACGAGCGGCCCC 983

Search completed: January 16, 2005, 09:30:23
Job time : 188.4 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 17:40:52 ; Search time 722.8 Seconds
(without alignments)
3925.542 Million cell updates/sec

Title: US-10-057-136-5

Perfect score: 60
Sequence: 1 ggtaccacgcgcgcgcctgc.....ccgaacacgcgcgcgcctccc 60

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	2297	6	BD272907 A recombi
2	36.6	61.0	2135	6	AX959916 Sequence
3	35.2	58.7	1320	6	E08764 cDNA encodi
4	35	58.3	120	6	AX192396 Sequence
5	35	58.3	120	6	BD000571 Human pol
6	35	58.3	180	6	E08763 cDNA encodi
7	35	58.3	1414	12	AF423031 Synthetic
8	35	58.3	1455	6	CQ715242 Sequence
9	35	58.3	1457	6	AX959914 Sequence
10	35	58.3	1721	6	CQ771290 Sequence
11	35	58.3	1721	6	AR492306 Sequence
12	35	58.3	1721	6	AX335860 Sequence
13	35	58.3	1721	6	AX440427 Sequence
14	35	58.3	1721	6	AX587588 Sequence
15	35	58.3	1721	9	HSTBYMA
16	35	58.3	1774	6	AX959684 Sequence
17	35	58.3	1774	6	AX959912 Sequence
18	35	58.3	1804	6	AR492318 Sequence
19	35	58.3	1804	6	AX335367 Sequence

20	35	58.3	1804	9	HUMMUCAB	J05581 Human polym
21	35	58.3	1834	12	AF423030	AF423030 Synthetic
22	35	58.3	1835	6	AX959918	AX959918 Sequence
23	35	58.3	1971	6	AX963157	AX963157 Sequence
24	35	58.3	2037	6	AX963159	AX963159 Sequence
25	35	58.3	2238	9	HSSETA	X52228 Human mRNA
26	35	58.3	4139	6	CQ834017	CQ834017 Sequence
27	35	58.3	4139	6	AX334899	AX334899 Sequence
28	35	58.3	4139	6	AX335372	AX335372 Sequence
29	35	58.3	4139	6	AX336712	AX336712 Sequence
30	35	58.3	4139	6	AX409474	AX409474 Sequence
31	35	58.3	4139	6	AX440481	AX440481 Sequence
32	35	58.3	4139	9	HUMPANMU	J05582 Human pancr
33	35	58.3	7188	9	AY463543	AY463543 Homo sapi
34	35	58.3	8181	6	AX406624	AX406624 Sequence
35	35	58.3	8181	9	HUMPEM	M61170 Human polym
36	35	58.3	8186	6	AR492320	AR492320 Sequence
37	35	58.3	133525	9	AL713999	AL713999 Human DNA
38	35	57.0	1258	9	AY261332	AY261332 Homo sapi
39	35	57.0	1935	9	AY216268	AY216268 Homo sapi
40	35	57.0	3420	9	BSM804095	AL832784 Homo sapi
41	35	57.0	3492	9	BC042651	BC042651 Homo sapi
42	35	57.0	3499	9	AF445801	AF445801 Homo sapi
43	35	57.0	107689	9	AL354681	AL354681 Human DNA
44	35	56.0	491	9	HUMEPISIB1	M32739 Human episi
45	35	56.0	518	6	AR492402	AR492402 Sequence

ALIGNMENTS

RESULT 1
BD272907
LOCUS
DEFINITION A recombinant vector expressing multiple constitutimulatory molecules
ACCESSION BD272907.1 GI:33082675
VERSION JP 2002531133-A/1.
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 2297)
AUTHORS Schlom,J., Hodge,J. and Panicali,D.
TITLE A recombinant vector expressing multiple constitutimulatory molecules
JOURNAL Patent: JP 2002531133-A 1 24-SEP-2002;
COMMENT THE UNITED STATES OF AMERICA, THERION BIOLOGICS CORP
OS Artificial Sequence
PN JP 2002531133-A/1
PD 24-SEP-2002
PF 12-NOV-1999 JP 2000586927
PR 09-DEC-1998 US 60/111582
PI JEFFREY SCHLOM, JAMES HODGE, DENNIS PANICALI
PC C12N15/02, A61K35/12, A61K35/74, A61K35/76, A61K38/00, A61K39/00,
PC A61K39/12,
PC A61K39/125, A61K39/21, A61K39/235, A61K39/245, A61K39/275, A61K39/
PC 29, A61K48/00,
PC A61P1/04, A61P31/04, A61P31/10, A61P31/12, A61P35/00, A61P37/02, PC
A61P37/06,
PC C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12N7/00, C12Q1/02, G01N33/
PC 53, C12N15/00,
PC C12N5/00, A61K37/02
CC Description of Artificial Sequence: VECTOR SEQUENCE FH Key
FT source 1. .2297
FT location/Qualifiers
FEATURES
source location/Qualifiers
1. .2297
/organism="Artificial Sequence".
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN

Query Match 100.0%; Score 60; DB 6; Length 2297;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCGCTGCGCAGGAGTGACGTGCGCGCCCGACACAGCGCGCGCTCCC 60
|||||
Db 466 GGATCCACCGCGCGCGCTGCGCAGGAGTGACGTGCGCGCCCGACACAGCGCGCGCTCCC 525

RESULT 2
AX959916 2135 bp DNA linear PAT 14-JAN-2004
LOCUS
DEFINITION Sequence 21 from Patent WO03100060.
ACCESSION AX959916
VERSION
KEYWORDS AX959916.1 GI:40880144
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Burden, N.G., Ellis, J.H. and Hamblin, P.A.
TITLE Muc-1 antigen with reduced number of vnter repeat units
JOURNAL Patent: WO 03100060-A 21 04-DEC-2003;
GLAXO GROUP LIMITED (GB)
FEATURES
source Location/Qualifiers
1..2135
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 61.0%; Score 36.6; DB 6; Length 2135;
Best Local Similarity 76.3%; Pred. No. 15;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCGCTGCGCAGGAGTGACGTGCGCGCCCGACACAGCGCGCTCC 59
|||||
Db 642 GGCTCCACCGCGCGCGCGCTGCGCAGGAGTGACGTGCGCGCCCGACACAGCGCGCTCC 700

RESULT 3
E08764 1320 bp RNA linear PAT 29-SEP-1997
LOCUS
DEFINITION cDNA encoding C-terminal fragment of human glycoprotein 39.
ACCESSION E08764
VERSION E08764.1 GI:2176876
KEYWORDS JP 1995051065-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1320)
Masuzawa, Y., Muramatsu, T. and Miyauchi, T.
GLYCOPROTEIN 39 GENE
Patent: JP 1995051065-A 2 28-FEB-1995;
NIPPON KOUTAI KENKYUSHO:KK, KAGOSHIMA UNIV
OS Homo sapiens (human)
PN JP 1995051065-A/2
PD 28-FEB-1995
PF 21-FEB-1992 JP 1992035085
PI MASUZAWA YASUSHI, MURAMATSU TAKASHI, MIYAUCHI TERUO PC
C12N15/09, G01N33/50//C07K14/00, C12P21/02;
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC Feature is identified by similarity;
FH Key Location/Qualifiers
FT source 1..1320
FT /organism='Homo sapiens'

FT /cell_type='gastric cancer'
FT /cell_line='KATO-3'
FT /clone_lib='lambda gt11 KATO-3 cDNA library'
FT /clone='lambda KP39',
FT mat_peptide 1..981
FT /product='C-terminal fragment of human FT
FT glycoprotein 39',
FT /partial
FT polyA_signal 1267..1272
FT polyA_site 1293..1320.
FEATURES
source Location/Qualifiers
1..1320
/organism="Homo sapiens"
/mol_type="genomic RNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 58.7%; Score 35.2; DB 6; Length 1320;
Best Local Similarity 76.8%; Pred. No. 38;
Matches 43; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCGCTGCGCAGGAGTGACGTGCGCGCCCGACACAGCGCGCGC 56
|||||
Db 1 GGCTCCACCGCGCGCGCGCTGCGCAGGAGTGACGTGCGCGCCCGACACAGCGCGCGC 56

RESULT 4
AX192396/c 120 bp DNA linear PAT 15-AUG-2001
LOCUS
DEFINITION Sequence 2 from Patent EP1103623.
ACCESSION AX192396
VERSION AX192396.1 GI:15210363
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Taylor-Papadimitriou, J., Burchell, J. and Gendler, S.
TITLE Human mucin core protein: nucleic acid probes, peptide fragments
and antibodies thereto, and uses thereof in diagnostic and
therapeutic methods
JOURNAL Patent: EP 1103623-A 2 30-MAY-2001;
IMPERIAL CANCER RESEARCH TECHNOLOGY LIMITED (GB)
FEATURES
source Location/Qualifiers
1..120
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 58.3%; Score 35; DB 6; Length 120;
Best Local Similarity 74.6%; Pred. No. 74;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCGCTGCGCAGGAGTGACGTGCGCGCCCGACACAGCGCGCTCC 59
|||||
Db 87 GGCTCCACCGCGCGCGCGCTGCGCAGGAGTGACGTGCGCGCCCGACACAGCGCGCTCC 29

RESULT 5
BD000571/c 120 bp DNA linear PAT 31-JAN-2002
LOCUS
DEFINITION Human polymorphic epithelial mucin core protein and nucleic acid
encoding the protein.
ACCESSION BD000571
VERSION BD000571.1 GI:18623684
KEYWORDS JP 2000333675-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 120)

AUTHORS Papadimitrov,J.T., Jendora,S. and Bachieru,J.
TITLE Human polymorphic epithelial mucin core protein and nucleic acid
JOURNAL encoding the protein
Patent: JP 200033675-A 2 05-DEC-2000;
IMPERIAL CANCER RESEARCH TECHNOLOGY LTD
OS Homo sapiens (human)
COMMENT PN JP 200033675-A/2
PD 05-DEC-2000
PF 26-APR-2000 JP 2000125724 8700279 PR
PR 07-JAN-1987 GB 8700269,07-JAN-1987 GB 8726172 PI JOYCE
22-APR-1987 US 041306,09-NOV-1987 GB
TAYLOR PAPADIMITROV,SANDRA JENDORA,JOY BACHIERU PC
C12N15/02,A61K38/00,A61K39/395,A61K39/395,A61K49/00,A61P35/00, PC
C07K14/47,
PC C07K16/44,C12N5/10,C12P21/08// (C12N15/02,C12R1:91), (C12N5/10,
PC C12R1:91),
PC C12N15/00,A61K37/02,C12N5/00, (C12N15/00,C12R1:91), (C12N5/00,
PC C12R1:91)
CC
FH Key Location/Qualifiers
FT source 1..120
/organism='Homo sapiens (human)'.
FEATURES
source 1..120
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 58.3%; Score 35; DB 6; Length 120;
Best Local Similarity 74.6%; Pred. No. 74;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 GGATCCACCGCGCCCTGCGCAGCAGTGCAGTGGCGCCCGCAGACAGCGCCCGCTCC 59
|||
Db 87 GGCTCCACCGCGCCCGCCAGCCAGCGGTGTCACTCGGCCCGGACACAGCGCGCC 29
RESULT 6
E08763 180 bp RNA linear PAT 29-SEP-1997
LOCUS cDNA encoding N-terminal fragment of human glycoprotein 39.
DEFINITION E08763
ACCESSION E08763
VERSION E08763.1 GI:2176875
KEYWORDS JP 1995051065-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
REFERENCE 1 (bases 1 to 180)
Masuzawa,Y., Muramatsu,T. and Miyauchi,T.
GLYCOPROTEIN 39 GENE
Patent: JP 1995051065-A 1 28-FEB-1995;
NIPPON KOUTAI KENKYUSHO:KK, KAGOSHIMA UNIV
COMMENT OS Homo sapiens (human)
PN JP 1995051065-A/1
PD 28-FEB-1995
PF 21-FEB-1992 JP 1992035085
PI MASUZAWA YASUSHI, MURAMATSU TAKASHI, MIYAUCHI TERUO PC
C12N15/09,G01N33/50//C07K14/00,C12P21/02;
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC Feature is identified by similarity;
FH Key Location/Qualifiers
FT source 1..180
/organism='Homo sapiens'
/cell_type='gastric cancer'
/cell_line='KATO-3'
/clone_1ib='lambda gtl1 KATO-3 cDNA library'
FT /clone='lambda KP39'

FT mat_peptide 1..180
FT /product='N-terminal fragment of human FT
glycoprotein 39'
FT /partial
FT repeat_region 1..180
FT repeat_unit 1..60.
FEATURES Location/Qualifiers
source 1..180
/organism="Homo sapiens"
/mol_type="genomic RNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 58.3%; Score 35; DB 6; Length 180;
Best Local Similarity 74.6%; Pred. No. 68;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 GGATCCACCGCGCCCTGCGCAGCAGTGCAGTGGCGCCCGCAGACAGCGCCCGCTCC 59
|||
Db 121 GGCTCCACCGCGCCCGCCAGCCAGCGGTGTCACTCGGCCCGGACACAGCGCGCC 179
RESULT 7
AF423031 1414 bp mRNA linear SYN 10-JUL-2003
LOCUS
DEFINITION AF423031
ACCESSION Synthetic construct Homo sapiens mucin variant MUC1-CT58 (MUC1)
VERSION mRNA, complete cds; alternatively spliced.
KEYWORDS AF423031
SOURCE AF423031.1 GI:19338621
ORGANISM
synthetic construct
synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 1414)
Hinojosa-Kurtzberg,A.M., Johansson,M.E., Madsen,C.S., Hansson,G.C.
and Gendler,S.J.
Novel MUC1 splice variants contribute to mucin overexpression in
CPTN-deficient mice
JOURNAL Am. J. Physiol. Gastrointest. Liver Physiol. 284 (5), G853-G862
(2003)
MEDLINE 22570517
PUBMED 12529261
REFERENCE 2 (bases 1 to 1414)
Hinojosa-Kurtzberg,A.M. and Gendler,S.J.
TITLE Direct Submision
JOURNAL Submitted (21-SEP-2001) Biochemistry, Mayo Clinic, 13400 E. Shea
Boulevard, Scottsdale, AZ 85259, USA
FEATURES Location/Qualifiers
source 1..1414
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/note="Homo sapiens gene in transgenic Mus musculus
C57BL/6; isolated from intestinal mucosa"
1..1414
/gene="MUC1"
1..1386
/gene="MUC1"
/note="alternatively spliced; contains exon 6b resulting
in variant carboxy-terminal domain; lacks sites for
beta-catenin and Grb2 interactions; derived from Homo
sapiens"
/codon_start=1
/transl_table=11
/product="mucin variant MUC1-CT58"
/protein_id="AA186735.1"
/db_xref="GI:19338622"
/translation="MTPGTQSPFFLLLLTLVTVVGSCHASSTPGGKETSATORSS
VPSSTENAVSMSTSVLSHSPGSGSSTTGGDVTTLATBPASGSAITWGDVTSVP
VTRPALGSTPPPAHDVTSAPDNKPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTS
APDNRPALGSTAPPVHNTVSASGSASASTLVHNGTSARATTPPAKSTPSPISPH
SDPTTLASHTKTDASSTHSTVPLTSSNHTSPOLSTGVSFPFLSFHISNLQFNS
SLEDPSIDYQELQDISEMFLQIYKGGFLGLSNIKFRPGSVVVLTLAFREGTIV

HDVETQFNQYKTEAASRYNLTISDVSVDPFPFSAOSGAGVPGNGIALLVLCVIVA
LAIVYIALAVCCQCRKNYQGLDIFPARDTYHPMSEYPTYHTHGRVPPSSTDRSPYE
KEMRVDRERLA"

ORIGIN

Query Match 58.3%; Score 35; DB 12; Length 1414;
Best Local Similarity 74.6%; Pred. No. 41;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY 1 GGATCCACCGCGCGCTGCGCAGGAGTACGTGGCGCCCGACACAGCGCCCGCTCC 59
|||
Db 385 GGCTCCACCGCGCGCCCGCAGCCAGGTGTCACTCGGCCCGGACACAGCGCGCCCG 443

RESULT 8

LOCUS CQ715242 1455 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 1176 from Patent WO02068579.
ACCESSION CQ715242
VERSION CQ715242.1 GI:42276099
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof

JOURNAL Patent: WO 02068579-A 1176 06-SEP-2002;
PE Corporation (NY) (US)

FEATURES
source 1..1455
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 58.3%; Score 35; DB 6; Length 1455;
Best Local Similarity 74.6%; Pred. No. 41;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY 1 GGATCCACCGCGCGCTGCGCAGGAGTACGTGGCGCCCGACACAGCGCCCGCTCC 59
|||
Db 412 GGCTCCACCGCGCGCCCGCAGCCAGGTGTCACTCGGCCCGGACACAGCGCGCGCC 470

RESULT 9

LOCUS AX959914 1457 bp DNA linear PAT 14-JAN-2004
DEFINITION Sequence 19 from Patent WO03100060.
ACCESSION AX959914
VERSION AX959914.1 GI:40880143
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 Burden, N.G., Ellis, J.H. and Hamblin, P.A.
TITLE Muc-1 antigen with reduced number of vnter repeat units
JOURNAL Patent: WO 03100060-A 19 04-DEC-2003;
GLAXO GROUP LIMITED (GB)

FEATURES
source 1..1457
Location/Qualifiers
1..1457
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 58.3%; Score 35; DB 6; Length 1457;
Best Local Similarity 74.6%; Pred. No. 41;

Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY 1 GGATCCACCGCGCGCTGCGCAGGAGTACGTGGCGCCCGACACAGCGCCCGCTCC 59
|||
Db 403 GGCTCCACCGCGCGCCCGCAGCCAGGTGTCACTCGGCCCGGACACAGCGCGCGCC 461

RESULT 10

LOCUS CQ771290 1721 bp DNA linear PAT 04-MAR-2004
DEFINITION Sequence 1 from Patent WO2004009632.
ACCESSION CQ771290
VERSION CQ771290.1 GI:45125397
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Goletz, S. and Karsten, U.
TITLE Method for the production of an immunostimulating mucin (muc1)
JOURNAL Patent: WO 2004009632-A 1 29-JAN-2004;
Nemod Immuntherapie AG (DE)

FEATURES
source 1..1721
Location/Qualifiers
1..1721
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 58.3%; Score 35; DB 6; Length 1721;
Best Local Similarity 74.6%; Pred. No. 39;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY 1 GGATCCACCGCGCGCGCTGCGCAGGAGTACGTGGCGCCCGACACAGCGCCCGCTCC 59
|||
Db 562 GGCTCCACCGCGCGCCCGCAGCCAGGTGTCACTCGGCCCGGACACAGCGCGCGCC 620

RESULT 11

LOCUS AR492306 1721 bp DNA linear PAT 15-MAY-2004
DEFINITION Sequence 3 from patent US 6716627.
ACCESSION AR492306
VERSION AR492306.1 GI:47260880
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1721)

AUTHORS Dobie, K.W.
TITLE Antisense modulation of mucin 1, transmembrane expression
JOURNAL Patent: US 6716627-A 3 06-APR-2004;
FEATURES
source 1..1721
Location/Qualifiers
1..1721
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 58.3%; Score 35; DB 6; Length 1721;
Best Local Similarity 74.6%; Pred. No. 39;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY 1 GGATCCACCGCGCGCGCTGCGCAGGAGTACGTGGCGCCCGACACAGCGCCCGCTCC 59
|||
Db 562 GGCTCCACCGCGCGCCCGCAGCCAGGTGTCACTCGGCCCGGACACAGCGCGCGCC 620

RESULT 12

LOCUS AX335860 1721 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 6369 from Patent WO0194629.
ACCESSION AX335860

VERSION AX335860.1 GI:18126579
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
TITLE Horrigan, S., Soppet, D.R. and Weaver, Z.
JOURNAL Cancer gene determination and therapeutic screening using signature
gene sets
Patent: WO 0194629-A 6369 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES Location/Qualifiers
source 1..1721
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 58.3%; Score 35; DB 6; Length 1721;
Best Local Similarity 74.6%; Pred. No. 39;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 GGATCCACCGCGCGCGCTGCGCAGAGTAGTCGTGGCGCCGACACAGCGCCCGCTCC 59
|||
Db 562 GGCTCCACCGCGCGCCGCGCAGCCACGCTGTCACTCGGCCCGGACACAGCGCGCCCG 620
RESULT 13
AX440427 1721 bp DNA linear PAT 28-JUN-2002
LOCUS AX440427
DEFINITION Sequence 280 from Patent WO0190154.
ACCESSION AX440427
VERSION AX440427.1 GI:21665237
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Xu, J., Mitcham, J.L., Harlocker, S.L., Dillon, D.C., Secrist, H.,
TITLE Lodes, M.J., Algate, P.A., Fling, S.P., Mannion, J., Benson, D.R. and
Carter, D.
JOURNAL Compositions and methods for the therapy and diagnosis of ovarian
Patent: WO 0190154-A 280 29-NOV-2001;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source 1..1721
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 58.3%; Score 35; DB 6; Length 1721;
Best Local Similarity 74.6%; Pred. No. 39;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 GGATCCACCGCGCGCGCTGCGCAGAGTAGTCGTGGCGCCGACACAGCGCCCGCTCC 59
|||
Db 562 GGCTCCACCGCGCGCCGCGCAGCCACGCTGTCACTCGGCCCGGACACAGCGCGCCCG 620
RESULT 14
AX587588 1721 bp DNA linear PAT 10-JAN-2003
LOCUS AX587588
DEFINITION Sequence 58 from Patent WO0246467.
ACCESSION AX587588
VERSION AX587588.1 GI:28212336
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct

REFERENCE 1
AUTHORS Bertucci, F., Houlgate, R., Birnbaum, D., Nguyen, C., Viens, P. and
TITLE Fert, V.
JOURNAL Gene expression profiling of primary breast carcinomas using arrays
Patent: WO 0246467-A 58 13-JUN-2002;
Ipsogen (FR)
FEATURES Location/Qualifiers
source 1..1721
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="primer"
misc_feature 1..1721
/note="mucin 1, transmembrane (MUC1) gene."
ORIGIN
Query Match 58.3%; Score 35; DB 6; Length 1721;
Best Local Similarity 74.6%; Pred. No. 39;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 GGATCCACCGCGCGCGCTGCGCAGAGTAGTCGTGGCGCCGACACAGCGCCCGCTCC 59
|||
Db 562 GGCTCCACCGCGCGCCGCGCAGCCACGCTGTCACTCGGCCCGGACACAGCGCGCCCG 620
RESULT 15
HSTBYMA 1721 bp mRNA linear PRI 14-AUG-1995
LOCUS HSTBYMA
DEFINITION Human mRNA for transmembrane epithelial tumour mucin antigen.
ACCESSION X52229
VERSION X52229.1 GI:37053
KEYWORDS transmembrane protein; tumor antigen.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Hareuveni, M., Tsarfaty, I., Zaretsky, J., Kotkes, P., Horev, J.,
TITLE Zrihan, S., Weiss, M., Green, S., Lathe, R., Keydar, I. and
Wreschner, D.H.
JOURNAL A transcribed gene, containing a variable number of tandem repeats,
Patent: Eur. J. Biochem. 189 (3), 475-486 (1990)
MEDLINE 90276414
PUBMED 2112460
REFERENCE 2 (bases 1 to 1721)
AUTHORS Wreschner, D.H., Hareuveni, M., Tsarfaty, I., Smorodinsky, N.,
TITLE Horov, J., Zaretsky, J., Kotkes, P., Weiss, M., Lathe, R., Dion, A.S. and
Keydar, I.
JOURNAL Human epithelial tumor antigen cDNA sequences. Differential
splicing may generate multiple protein forms
Eur. J. Biochem. 189 (3), 463-473 (1990)
MEDLINE 90276413
PUBMED 2351132
REFERENCE 3 (bases 1 to 1721)
AUTHORS Wreschner, D.H.
TITLE Direct Submission
JOURNAL Submitted (16-MAR-1990) Wreschner D.H., Tel Aviv University, Dept
of Microbiology, Faculty of Life Sciences, Tel Aviv 69978, Israel
COMMENT See <X52228> for secreted tumour antigen.
FEATURES Location/Qualifiers
source 1..1721
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="1q21-24"
/cell_line="T47D breast tumour"
/clone_id="lambda gt11"
58..1605
CDS

```
/codon_start=1
/product="precursor polypeptide (AA -21 to 494)"
/protein_id="CA36478.1"
/db_xref="GI:37054"
/db_xref="GOA:P15941"
/db_xref="Swiss-Prot:P15941"
/translation="MPTGQSPFFLLLLTLVTVTSGHASTPGKEKETSATORSS
VPSSTEKNVAVSMSTSYLSSHPGSGSSTQGQDVTLPATEPAGSAATWGQDVTSP
VTRPALGSTTPPAHDVTSAPDNKPAPGSTAPPAHGVTSAPDTRPPGSTAPPAHGVTS
APDTRPPGSTAPPAHGVTSAPDTRAPAGSTAPPAHGVTSAPDNRPALASTAPVHN
TSAGSGASGSASTLVHNGTSARATTPASKSTPFISSHSDPTTLASHSTKDASS
THSTVPEPLTSSNHSSTPOLSTGVFFELSFHISNLQFNSLEDPTDYQGEIQDIS
EMFLQIKYQGGFLGSLNIKFRPGSVVQLTLAFREGTINVHDEVTQFNQKTEASRY
NLTISDVSVSDVPEPFSQAQAGVPGWGIALLVLCVLAIVYLLIALAVCCRRKN
YGQLDIFPARDTYHPNSEYPTYHTHGRYVPPSSTDRSPYEKVSAGNGSSLSYTNPAV
AATSANL"
sig_peptide      58..120
                  /note="signal peptide (AA -21 to -1)"
mat_peptide      121..1602
                  /product="transmembrane tumor mucin antigen (AA 1 to 494)"
ORIGIN
```

```
Query Match      58.3%; Score 35; DB 9; Length 1721;
Best Local Similarity 74.6%; Pred. No. 39;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY      1 GGATCCACCGCGCGCTCGCGACGAGTGACGTGGCGCGCGACACGCGCGCTCC 59
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      562 GGCTCCACCGCGCGCGCGACGCCACGATGTCACCTCGCGCGCGACACGCGCGCTCC 620
```

Search completed: January 15, 2005, 22:36:46
Job time : 723.8 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 19:50:13 ; Search time 35.2 Seconds
(without alignments)
1211.572 Million cell updates/sec

Title: US-10-057-136-4

Perfect score: 60
Sequence: 1 ggcagctactgcaccacccggc.....ctgatcacagacctgcacct 60

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 segs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCITUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfileseq1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	27.4	45.7	984	4	US-09-543-681A-345
C 2	27.4	45.0	1804	4	US-10-029-517-17
C 3	27.4	45.0	8186	4	US-10-029-517-19
C 4	26.4	44.0	3722	4	US-09-016-434-1129
C 5	26.4	43.3	2210	3	US-08-464-700-1
C 6	25.8	43.0	489	4	US-09-270-767-10939
C 7	25.6	42.7	518	4	US-10-029-517-101
C 8	25.6	42.7	981	4	US-10-029-517-16
C 9	25.6	42.7	1721	4	US-10-029-517-3
C 10	25.6	42.7	3343	4	US-10-029-517-102
C 11	25.4	42.3	572	4	US-10-029-517-18
C 12	25.4	41.7	1260	4	US-09-328-352-2989
C 13	24.8	41.3	1351	3	US-09-175-928-41
C 14	24.8	41.3	1362	4	US-09-489-039A-3117
C 15	24.8	41.3	229354	4	US-09-765-400-64
C 16	24.8	41.3	229354	4	US-09-765-400-64
C 17	24.6	41.0	318	4	US-09-489-039A-3181
C 18	24.6	41.0	798	4	US-09-489-039A-3222
C 19	24.4	40.7	2159	3	US-08-286-870A-7
C 20	24.4	40.3	60	4	US-09-475-947A-246
C 21	24.2	40.3	1599	4	US-09-475-515-54
C 22	24.2	40.3	2031	4	US-09-475-515-60
C 23	24.2	40.3	2112	4	US-09-475-515-56
C 24	24.2	40.3	2112	4	US-09-475-515-57
C 25	24.2	40.3	2181	4	US-09-475-515-58
C 26	24.2	40.3	2553	4	US-09-475-515-66
C 27	24.2	40.3	2634	4	US-09-475-515-64

28	24.2	40.3	4766	4	US-09-475-515-73	Sequence 73, Appl
29	23.8	39.7	400	4	US-09-401-064-341	Sequence 341, App
C 30	23.8	39.7	463	4	US-09-841-334A-7	Sequence 7, Appli
C 31	23.8	39.7	463	4	US-09-837-969A-7	Sequence 25, Appl
C 32	23.8	39.7	1298	4	US-09-311-784A-25	Sequence 280, App
C 33	23.8	39.7	2292	4	US-08-073-807A-1	Sequence 1, Appli
C 34	23.8	39.7	2455	1	US-08-023-655-1003	Sequence 1003, Ap
C 35	23.8	39.7	4020	3	US-09-050-159-130	Sequence 5, Appli
C 36	23.8	39.7	4020	4	US-09-023-655-1003	Sequence 42, Appl
C 37	23.6	39.3	165	1	US-08-458-298-5	Sequence 5, Appli
C 38	23.6	39.3	165	1	US-08-458-298-5	Sequence 46, Appl
C 39	23.6	39.3	519	4	US-09-646-028-42	Sequence 1284, Ap
C 40	23.6	39.3	534	4	US-09-270-767-1284	Sequence 16566, A
C 41	23.6	39.3	789	4	US-09-270-767-16566	Sequence 1, Appli
C 42	23.6	39.3	1833	4	US-09-129-603-1	Sequence 1027, Ap
C 43	23.6	39.3	1833	4	US-09-129-603-1	
C 44	23.6	39.3	5399	4	US-08-998-416-1027	
C 45	23.4	39.0	733	3	US-08-998-416-1027	

ALIGNMENTS

RESULT 1
US-09-543-681A-345/c
; Sequence 345, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 345
; LENGTH: 984
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-345

Query Match 45.7%; Score 27.4; DB 4; Length 984;
Best Local Similarity 69.8%; Pred. No. 1.1;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 6 TACTGACCCACCGGCATGCGTACATCAGCAGCTGATCAAGACCTGCAC 58
Db 546 TAATACACCTCCGCCACCTTCCATACCTTCATACCTTAAGACAAATGTCAC 494

RESULT 2
US-10-029-517-17
; Sequence 17, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 17
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (73)...(1500)
US-10-029-517-17

Query Match 45.0%; Score 27; DB 4; Length 1804;
Best Local Similarity 66.1%; Pred. No. 1.8;
Matches 39; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
QY 1 GGCAGTACTGCACCCGCGACATGCGCTAACATCAGCAGCCTGATACAGACCTGCACC 59
DB 457 GGCTCCACCGCCCCCGCAGCCGCGTGTACCTCGGCCCGGACACCAAGCGCGCCCC 515

RESULT 3
US-10-029-517-19
; Sequence 19, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 19
; LENGTH: 8186
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 6899
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7155
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7184
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7957
; OTHER INFORMATION: unknown
; NAME/KEY: unknown
; LOCATION: (2997)...(3498)
; OTHER INFORMATION: Intron 1
; NAME/KEY: Intron:exon junction
; LOCATION: (3498)...(3499)
; OTHER INFORMATION: Intron 1:exon 2
; NAME/KEY: exon
; LOCATION: (3508)...(3599)
; OTHER INFORMATION: exon 2d
; NAME/KEY: exon:intron junction
; LOCATION: (3982)...(3983)
; OTHER INFORMATION: exon 2a:intron 2a
; NAME/KEY: Intron:exon junction
; LOCATION: (4205)...(4206)
; OTHER INFORMATION: Intron 2c:exon 3c
; NAME/KEY: Intron:exon junction
; LOCATION: (4259)...(4260)
; OTHER INFORMATION: Intron 2d:exon 3d
; NAME/KEY: exon
; LOCATION: (4260)...(4328)
; OTHER INFORMATION: exon 3d
; NAME/KEY: Intron:exon junction
; LOCATION: (4632)...(4633)
; OTHER INFORMATION: Intron 3:exon 4
; NAME/KEY: exon
; LOCATION: (4914)...(5035)
; OTHER INFORMATION: exon 5
; NAME/KEY: Intron
; LOCATION: (5266)...(6293)
; OTHER INFORMATION: Intron 6
US-10-029-517-19

Query Match 45.0%; Score 27; DB 4; Length 8186;
Best Local Similarity 66.1%; Pred. No. 2.6;
Matches 39; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 GGCAGTACTGCACCCGCGACATGCGCTAACATCAGCAGCCTGATACAGACCTGCACC 59
DB 3825 GGCTCCACCGCCCCCGCAGCCGCGTGTACCTCGGCCCGGACACCAAGCGCGCCCC 3883

RESULT 4
US-09-016-434-1129/c
; Sequence 1129, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1129:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3722 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g1399211
US-09-016-434-1129

Query Match 44.0%; Score 26.4; DB 4; Length 3722;
Best Local Similarity 65.0%; Pred. No. 3.5;
Matches 39; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 GGCAGTACTGCACCCGCGACATGCGCTAACATCAGCAGCCTGATACAGACCTGCACC 60
DB 2384 GGTGGCCAGCGCCCTCCCGACCTTGTGTGCCACCACTGGCAGAGCCCTTCACT 2325

RESULT 5
US-08-464-700-1/c
; Sequence 1, Application US/08464700
; Patent No. 6232458
; GENERAL INFORMATION:
; APPLICANT: WEISS, ANTHONY S
; APPLICANT: MARTIN, STEPHEN L
; TITLE OF INVENTION: SYNTHETIC POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:

```

; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, PO Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,700
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PL6520
; FILING DATE: 22-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PL9661
; FILING DATE: 28-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU93/00655
; FILING DATE: 16-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GHC3USA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2210 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; US-08-464-700-1

Query Match          43.3%; Score 26; DB 3; Length 2210;
Best Local Similarity 70.0%; Pred. No. 4.3;
Matches 35; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY      10 GCACCAACCGGACATGCGGTACATCAGCACCCTGATACAGACCTGCACC 59
Db      1459 GCAACACCTACACCTGCTGTGCAACGCCAACCCTGCTGACCTACGCC 1410

RESULT 6
US-09-270-767-10939
; Sequence 10939, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 10939
; LENGTH: 489
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: n means any nucleotide
US-09-270-767-10939

Query Match          43.0%; Score 25.8; DB 4; Length 489;
Best Local Similarity 67.9%; Pred. No. 3.4;
Matches 36; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
```

```

QY      8 CTGACCAACCGGACATGCGGTACATCAGCACCCTGATACAGACCTGCACC 60
Db      304 CAGCAATCATCAGCTGTGCAACATCAGAGCTGCAGCAACATCAGACCT 356

RESULT 7
US-10-029-517-101
; Sequence 101, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 101
; LENGTH: 518
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; US-10-029-517-101

Query Match          42.7%; Score 25.6; DB 4; Length 518;
Best Local Similarity 66.1%; Pred. No. 4;
Matches 37; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY      1 GGCAGTACTGCACCAACCGGACATGCGGTACATCAGCACCCTGATACAGACCTGC 56
Db      462 GGCTCCACCGCCCCCGGACCGGTGTACCTCGGCCCCGGACCAAGGCCGCG 517

RESULT 8
US-10-029-517-16
; Sequence 16, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 16
; LENGTH: 981
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon:exon junction
; LOCATION: (464)...(465)
; OTHER INFORMATION: exon 3b:exon 4
US-10-029-517-16

Query Match          42.7%; Score 25.6; DB 4; Length 981;
Best Local Similarity 66.1%; Pred. No. 4.8;
Matches 37; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY      1 GGCAGTACTGCACCAACCGGACATGCGGTACATCAGCACCCTGATACAGACCTGC 56
Db      21 GGCTCCACCGCCCCCGGACCGGTGTACCTCGGCCCCGGACCAAGGCCGCG 76

RESULT 9
US-10-029-517-3
; Sequence 3, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
```

```
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 3
; LENGTH: 1721
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (58) ... (1605)
US-10-029-517-3
```

```
Query Match      42.7%; Score 25.6; DB 4; Length 1721;
Best Local Similarity 66.1%; Pred. No. 5.5;
Matches 37; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
```

```
OY      1 GGCAGTACTGCACCGGACCATGGCGTAACATCAGCACCTGTATACAGACCTGC 56
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      622 GGCTCCACCGCCCCCGCCAGCCCATGTGTCACTCGGCCCCGGACAACAGGCCCGC 677
```

```
RESULT 10
US-10-029-517-102
; Sequence 102, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 102
; LENGTH: 3343
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-10-029-517-102
```

```
Query Match      42.7%; Score 25.6; DB 4; Length 3343;
Best Local Similarity 66.1%; Pred. No. 6.6;
Matches 37; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
```

```
OY      1 GGCAGTACTGCACCGGACCATGGCGTAACATCAGCACCTGTATACAGACCTGC 56
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1728 GGCTCCACCGCCCCCGCCAGCCCATGTGTCACTCGGCCCCGGACAACAGGCCCGC 1783
```

```
RESULT 11
US-10-029-517-18
; Sequence 18, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 18
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (67) ... (572)
US-10-029-517-18
```

```
Query Match      42.3%; Score 25.4; DB 4; Length 572;
Best Local Similarity 64.4%; Pred. No. 4.9;
```

```
Matches 38; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
OY      1 GGCAGTACTGCACCGGACCATGGCGTAACATCAGCACCTGTATACAGACCTGCACC 59
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      478 GGCTCCACCGCCCCCGCCAGCCCATGTGTCACTCGGCCCCGGACAACAGGCCCGC 536
```

```
RESULT 12
US-09-328-352-2989/C
; Sequence 2989, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 2989
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-2989
```

```
Query Match      41.7%; Score 25; DB 4; Length 1260;
Best Local Similarity 64.9%; Pred. No. 8.4;
Matches 37; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
```

```
OY      4 AGTACTGCACCGGACCATGGCGTAACATCAGCACCTGTATACAGACCTGCACCT 60
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      470 AGAATTTCACCCCATGCACTGCTGTACCATGACCACCGGTTAAAGTAATTGAACCT 414
```

```
RESULT 13
US-09-175-928-41
; Sequence 41, Application US/09175928A
; Patent No. 6312921
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallee, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Mi, Sha
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6006B.AJ172A
; CURRENT APPLICATION NUMBER: US/09/175,928A
; CURRENT FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 41
; LENGTH: 1351
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (134)
US-09-175-928-41
```

```
Query Match      41.3%; Score 24.8; DB 3; Length 1351;
Best Local Similarity 67.3%; Pred. No. 10;
Matches 35; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
```

```
OY      7 ACTGCACCGGACCATGGCGTAACATCAGCACCTGTATACAGACCTGCAC 58
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      825 ACTGTATCCCGCACATACAGCAACGCCACGCTCTGTATCAAGATCTTCAC 876
```

```
RESULT 14
```

```

US-09-489-039A-3117/c
; Sequence 3117, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709,2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 3117
; LENGTH: 1362
;
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-3117

```


This Page Blank (uspto)


```
Location/Qualifiers
1..370
/organism="Bombyx mori"
/mol_type="mRNA"
/db_xref="taxon:7091"
clone="tess0695"
/tissue_type="testis"
/dev_stage="spinning stage"
```

ORIGIN	/clone_lib="test"
Query Match	48.7%; Score 29.2; DB 5; Length 370;
Best Local Similarity	74.0%; Pred. No. 22;
Matches 37; Conservative	0; Mismatches 13; Indels 0; Gaps 0;
QY	11 CACCACCGGCACATGGCGTAACTAGACACCTGATACAGACCTGCACCT 60
Db	274 CAGCACCGGCACCTGAACACGACACGACACCTGAACACGACCGGCACCT 323
RESULT 8	
CK531267/c	
LOCUS	CK531267 375 bp mRNA linear EST 15-JAN-2004
DEFINITION	rswhao 009036.y1 swg Bombyx mori cDNA, mRNA sequence.
ACCESSION	CK531267
VERSION	CK531267.1 GI:40915720
KEYWORDS	EST.
SOURCE	Bombyx mori (domestic silkworm)
ORGANISM	Bombyx mori
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea; Bombycidae; Bombyx.
REFERENCE	1 (bases 1 to 375)
AUTHORS	Xia,Q., Zhou,Z., Lu,C., Cheng,D., Dai,F., Li,B., Zhao,P., Zha,X., Cheng,T., Chai,C., Pan,G., Xu,J., Liu,C., Lin,Y., Qian,J., Hou,Y., Wu,Z., Li,G., Pan,M., Li,C., Shen,Y., Lan,X., Yuan,L., Yang,G., Wan,Y., Zhu,Y., Yu,M., Shen,W., Wu,D., Xiang,Z., Yu,J., Wang,J., Li,R., Li,H., Li,G., Su,J., Wang,X., Li,G., Zhang,Z., Li,J., Zhang,Q., Wei,N., Xu,J., Sun,H., Dong,L., Liu,D., Zhao,S., Zhao,X., Meng,Q., Lan,F., Fang,L., Li,Y., Wang,J., Ye,J., Wong,G.K.-S., Yang,H., Shi,J., Wu,Q., Li,C., Li,D., Sun,Y., Zhang,Z., Yang,Z., Huang,Y., Xi,Y., Qi,Q., He,D., Huang,H., Zhang,X., Wang,Z., Li,W. and Cao,Y.
TITLE	A draft sequence for the genome of the domesticated silkworm (Bombyx mori)
JOURNAL	Unpublished (2004)
COMMENT	Contact: Yang Huan Ming Beijing Genomics Institute Chinese Academy of Sciences Beijing Airport Industrial Zone B6, Beijing 101300, P.R.China Email: hyang@igtp.ac.cn.
FEATURES	Location/Qualifiers
source	1..375
	/organism="Bombyx mori"
	/mol_type="mRNA"
	/strain="Dazhao (P50)"
	/db_xref="taxon:7091"
	/sex="male"
	/tissue_type="testis"
	/dev_stage="5th-instar day-3 larva"
	/clone_lib="swg"
	/note="Vector: pBluescript II SK(+)"
ORIGIN	
Query Match	48.7%; Score 29.2; DB 7; Length 375;
Best Local Similarity	74.0%; Pred. No. 22;
Matches 37; Conservative	0; Mismatches 13; Indels 0; Gaps 0;
QY	11 CACCACCGGCACATGGCGTAACTAGACACCTGATACAGACCTGCACCT 60
Db	267 CAGCACCGGCACCTGAACACGACACGACACCTGAACACGACCGGCACCT 218
RESULT 9	
CK542714	
LOCUS	CK542714 375 bp mRNA linear EST 15-JAN-2004
DEFINITION	rswhb0 008394.y1 swn Bombyx mori cDNA, mRNA sequence.
ACCESSION	CK542714
VERSION	CK542714.1 GI:40927168
KEYWORDS	EST.
SOURCE	Bombyx mori (domestic silkworm)

ORGANISM Bombyx mori
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea; Bombycidae; Bombyx.

REFERENCE
AUTHORS
Xia,Q., Zhou,Z., Lu,C., Cheng,D., Dai,F., Li,B., Zhao,P., Zha,X., Cheng,T., Chai,C., Pan,G., Xu,J., Liu,C., Lin,Y., Qian,J., Hou,Y., Wu,Z., Li,G., Pan,M., Li,C., Shen,Y., Lan,X., Yuan,L., Yang,G., Wan,Y., Zhu,Y., Yu,M., Shen,W., Wu,D., Xiang,Z., Yu,J., Wang,J., Li,R., Li,H., Li,G., Su,J., Wang,X., Li,G., Zhang,Z., Li,J., Zhang,Q., Wei,N., Xu,J., Sun,H., Dong,L., Liu,D., Zhao,S., Zhao,X., Meng,Q., Ian,F., Fang,L., Li,Y., Wang,J., Ye,J., Wong,G.K.-S., Yang,H., Shi,J., Wu,Q., Li,C., Li,D., Sun,Y., Zhang,Z., Yang,Z., Huang,Y., Xi,Y., Qi,Q., He,D., Huang,H., Zhang,X., Wang,Z., Li,W. and Cao,Y.
A draft sequence for the genome of the domesticated silkworm (Bombyx mori)
Unpublished (2004)
Contact: Yang Huan Ming
Beijing Genomics Institute
Chinese Academy of Sciences
Beijing Airport Industrial Zone B6, Beijing 101300, P.R.China
Email: hyang@igtp.ac.cn.
location/Qualifiers
1..375
/organism="Bombyx mori"
/mol_type="mRNA"
/strain="Dazhao(P50)"
/db_xref="taxon:7091"
/sex="female"
/tissue_type="ovary"
/dev_stage="5th-instar day-3 larva"
/clone_lib="swl"
/note="Vector: pBluescript II SK(+)"

ORIGIN

Query Match	48.7%;	Score 29.2;	DB 7;	Length 375;
Best Local Similarity	74.0%;	Pred. No. 22;	Mismatches 13;	Indels 0;
Matches	37;	Conservative 0;	Gaps 0;	
Dn	278	CAGCACCGGACCTGAACCAGCAGCAGCACTGAACCAGCACCGGCACCT	327	
Qy	11	CACCAACGGGACATGGCGTAACTCAGCACCTGTACAAGACTGCACCT	60	

RESULT 10
CK487724
LOCUS CK487724
DEFINITION rswab0_005006.y1 swa Bombyx mori cDNA, mRNA sequence.
ACCESSION CK487724
VERSION CK487724.1 GI:40843303
KEYWORDS EST.
SOURCE Bombyx mori (domestic silkworm)
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea; Bombycidae; Bombyx.

REFERENCE
AUTHORS
Xia,Q., Zhou,Z., Lu,C., Cheng,D., Dai,F., Li,B., Zhao,P., Zha,X., Cheng,T., Chai,C., Pan,G., Xu,J., Liu,C., Lin,Y., Qian,J., Hou,Y., Wu,Z., Li,G., Pan,M., Li,C., Shen,Y., Lan,X., Yuan,L., Yang,G., Wan,Y., Zhu,Y., Yu,M., Shen,W., Wu,D., Xiang,Z., Yu,J., Wang,J., Li,R., Li,H., Li,G., Su,J., Wang,X., Li,G., Zhang,Z., Li,J., Zhang,Q., Wei,N., Xu,J., Sun,H., Dong,L., Liu,D., Zhao,S., Zhao,X., Meng,Q., Ian,F., Fang,L., Li,Y., Wang,J., Ye,J., Wong,G.K.-S., Yang,H., Shi,J., Wu,Q., Li,C., Li,D., Sun,Y., Zhang,Z., Yang,Z., Huang,Y., Xi,Y., Qi,Q., He,D., Huang,H., Zhang,X., Wang,Z., Li,W. and Cao,Y.
A draft sequence for the genome of the domesticated silkworm (Bombyx mori)
Unpublished (2004)
Contact: Yang Huan Ming
Beijing Genomics Institute
Chinese Academy of Sciences
Beijing Airport Industrial Zone B6, Beijing 101300, P.R.China
Email: hyang@igtp.ac.cn.
location/Qualifiers
1..375
/organism="Bombyx mori"
/mol_type="mRNA"
/strain="Dazhao(P50)"
/db_xref="taxon:7091"
/sex="female"
/tissue_type="ovary"
/dev_stage="5th-instar day-3 larva"
/clone_lib="swl"
/note="Vector: pBluescript II SK(+)"

TITLE
JOURNAL COMMENT
Unpublished (2004)

FEATURES
source
1..375
/organism="Bombyx mori"
/mol_type="mRNA"
/strain="Dazhao(P50)"
/db_xref="taxon:7091"
/sex="female"
/tissue_type="ovary"
/dev_stage="5th-instar day-3 larva"
/clone_lib="swl"
/note="Vector: pBluescript II SK(+)"

JOURNAL COMMENT
Unpublished (2004)
Contact: Yang Huan Ming
Beijing Genomics Institute
Chinese Academy of Sciences
Beijing Airport Industrial Zone B6, Beijing 101300, P.R.China
Email: hyang@igtp.ac.cn.
location/Qualifiers
1..375
/organism="Bombyx mori"
/mol_type="mRNA"
/strain="Dazhao(P50)"
/db_xref="taxon:7091"
/sex="female"
/tissue_type="ovary"
/dev_stage="5th-instar day-3 larva"
/clone_lib="swl"
/note="Vector: pBluescript II SK(+)"

TITLE
JOURNAL COMMENT
Unpublished (2004)
Contact: Yang Huan Ming
Beijing Genomics Institute
Chinese Academy of Sciences
Beijing Airport Industrial Zone B6, Beijing 101300, P.R.China
Email: hyang@igtp.ac.cn.
location/Qualifiers
1..375
/organism="Bombyx mori"
/mol_type="mRNA"
/strain="Dazhao(P50)"
/db_xref="taxon:7091"
/sex="female"
/tissue_type="ovary"
/dev_stage="5th-instar day-3 larva"
/clone_lib="swl"
/note="Vector: pBluescript II SK(+)"

TITLE
JOURNAL COMMENT
Unpublished (2004)
Contact: Yang Huan Ming
Beijing Genomics Institute
Chinese Academy of Sciences
Beijing Airport Industrial Zone B6, Beijing 101300, P.R.China
Email: hyang@igtp.ac.cn.
location/Qualifiers
1..375
/organism="Bombyx mori"
/mol_type="mRNA"
/strain="Dazhao(P50)"
/db_xref="taxon:7091"
/sex="female"
/tissue_type="ovary"
/dev_stage="5th-instar day-3 larva"
/clone_lib="swl"
/note="Vector: pBluescript II SK(+)"

TITLE
JOURNAL COMMENT
Unpublished (2004)
Contact: Yang Huan Ming
Beijing Genomics Institute
Chinese Academy of Sciences
Beijing Airport Industrial Zone B6, Beijing 101300, P.R.China
Email: hyang@igtp.ac.cn.
location/Qualifiers
1..375
/organism="Bombyx mori"
/mol_type="mRNA"
/strain="Dazhao(P50)"
/db_xref="taxon:7091"
/sex="female"
/tissue_type="ovary"
/dev_stage="5th-instar day-3 larva"
/clone_lib="swl"
/note="Vector: pBluescript II SK(+)"

TITLE
JOURNAL COMMENT
Unpublished (2004)
Contact: Yang Huan Ming
Beijing Genomics Institute
Chinese Academy of Sciences
Beijing Airport Industrial Zone B6, Beijing 101300, P.R.China
Email: hyang@igtp.ac.cn.
location/Qualifiers
1..375
/organism="Bombyx mori"
/mol_type="mRNA"
/strain="Dazhao(P50)"
/db_xref="taxon:7091"
/sex="female"
/tissue_type="ovary"
/dev_stage="5th-instar day-3 larva"
/clone_lib="swl"
/note="Vector: pBluescript II SK(+)"

TITLE
JOURNAL COMMENT
Unpublished (2004)
Contact: Yang Huan Ming
Beijing Genomics Institute
Chinese Academy of Sciences
Beijing Airport Industrial Zone B6, Beijing 101300, P.R.China
Email: hyang@igtp.ac.cn.
location/Qualifiers
1..375
/organism="Bombyx mori"
/mol_type="mRNA"
/strain="Dazhao(P50)"
/db_xref="taxon:7091"
/sex="female"
/tissue_type="ovary"
/dev_stage="5th-instar day-3 larva"
/clone_lib="swl"
/note="Vector: pBluescript II SK(+)"

TITLE
JOURNAL COMMENT
Unpublished (2004)
Contact: Yang Huan Ming
Beijing Genomics Institute
Chinese Academy of Sciences
Beijing Airport Industrial Zone B6, Beijing 101300, P.R.China
Email: hyang@igtp.ac.cn.
location/Qualifiers
1..375
/organism="Bombyx mori"
/mol_type="mRNA"
/strain="Dazhao(P50)"
/db_xref="taxon:7091"
/sex="female"
/tissue_type="ovary"
/dev_stage="5th-instar day-3 larva"
/clone_lib="swl"
/note="Vector: pBluescript II SK(+)"

TITLE
JOURNAL COMMENT
Unpublished (2004)
Contact: Yang Huan Ming
Beijing Genomics Institute
Chinese Academy of Sciences
Beijing Airport Industrial Zone B6, Beijing 101300, P.R.China
Email: hyang@igtp.ac.cn.
location/Qualifiers
1..375
/organism="Bombyx mori"
/mol_type="mRNA"
/strain="Dazhao(P50)"
/db_xref="taxon:7091"
/sex="female"
/tissue_type="ovary"
/dev_stage="5th-instar day-3 larva"
/clone_lib="swl"
/note="Vector: pBluescript II SK(+)"

TITLE
JOURNAL COMMENT
Unpublished (2004)
Contact: Yang Huan Ming
Beijing Genomics Institute
Chinese Academy of Sciences
Beijing Airport Industrial Zone B6, Beijing 101300, P.R.China
Email: hyang@igtp.ac.cn.
location/Qualifiers
1..375
/organism="Bombyx mori"
/mol_type="mRNA"
/strain="Dazhao(P50)"
/db_xref="taxon:7091"
/sex="female"
/tissue_type="ovary"
/dev_stage="5th-instar day-3 larva"
/clone_lib="swl"
/note="Vector: pBluescript II SK(+)"

TITLE
JOURNAL COMMENT
Unpublished (2004)
Contact: Yang Huan Ming
Beijing Genomics Institute
Chinese Academy of Sciences
Beijing Airport Industrial Zone B6, Beijing 101300, P.R.China
Email: hyang@igtp.ac.cn.
location/Qualifiers
1..375
/organism="Bombyx mori"
/mol_type="mRNA"
/strain="Dazhao(P50)"
/db_xref="taxon:7091"
/sex="female"
/tissue_type="ovary"
/dev_stage="5th-instar day-3 larva"
/clone_lib="swl"
/note="Vector: pBluescript II SK(+)"

TITLE
JOURNAL COMMENT
Unpublished (2004)
Contact: Yang Huan Ming
Beijing Genomics Institute
Chinese Academy of Sciences
Beijing Airport Industrial Zone B6, Beijing 101300, P.R.China
Email: hyang@igtp.ac.cn.
location/Qualifiers
1..375
/organism="Bombyx mori"
/mol_type="mRNA"
/strain="Dazhao(P50)"
/db_xref="taxon:7091"
/sex="female"
/tissue_type="ovary"
/dev_stage="5th-instar day-3 larva"
/clone_lib="swl"
/note="Vector: pBluescript II SK(+)"

TITLE
JOURNAL COMMENT
Unpublished (2004)
Contact: Yang Huan Ming
Beijing Genomics Institute
Chinese Academy of Sciences
Beijing Airport Industrial Zone B6, Beijing 101300, P.R.China
Email: hyang@igtp.ac.cn.
location/Qualifiers
1..375
/organism="Bombyx mori"
/mol_type="mRNA"
/strain="Dazhao(P50)"
/db_xref="taxon:7091"
/sex="female"
/tissue_type="ovary"
/dev_stage="5th-instar day-3 lar

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 16:00:11 ; Search time 170.5 Seconds
(without alignments)
1847.304 Million cell updates/sec

Title: US-10-057-136-4
Perfect score: 60
Sequence: 1 ggcagctactgcaccaccgcgc.....ctgatacaagaccgcacact 60

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0.

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : N_Geneseq_23Sep04:*
- 1: geneseqn1980s:*
 - 2: geneseqn1990s:*
 - 3: geneseqn2000s:*
 - 4: geneseqn2001as:*
 - 5: geneseqn2001bs:*
 - 6: geneseqn2002as:*
 - 7: geneseqn2002bs:*
 - 8: geneseqn2003as:*
 - 9: geneseqn2003bs:*
 - 10: geneseqn2003cs:*
 - 11: geneseqn2003ds:*
 - 12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	60	100.0	60	2	AAV48317	Aav48317 Nucleotid
2	37.6	62.7	60	2	AAV48325	Aav48325 Nucleotid
3	36.6	61.0	525	3	AAD00385	Aad00385 Human Muc
4	36.6	61.0	891	3	AAD00391	Aad00391 Ubiquitin
5	35.4	59.0	60	2	AAV48323	Aav48323 Nucleotid
6	35	58.3	1371	3	AAD00388	Aad00388 Human Muc
7	35	58.3	1737	3	AAD00394	Aad00394 Ubiquitin
8	33.4	55.7	60	2	AAV48321	Aav48321 Nucleotid
9	33.4	55.7	60	2	AAV48322	Aav48322 Nucleotid
10	33.2	55.3	60	2	AAV48320	Aav48320 Nucleotid
11	29	48.3	60	2	AAV48324	Aav48324 Nucleotid
12	28.6	47.7	110000	6	ABA90521	Continuation (8 of
13	28.2	47.0	4144	12	AD157666	Ad157666 Human bre
14	28	46.7	584	8	ACC55474	Acc55474 Rice endo
15	28	46.7	717	10	ADG14257	Adg14257 Cellulose
16	28	46.7	5562	2	AAT86625	Aat86625 C. thermo
17	27.4	45.7	60	2	AAV48319	Aav48319 Nucleotid
18	27.4	45.7	984	10	ADF00060	Adf00060 Bacterial
19	27.4	45.7	1682	8	AAL51696	Aal51696 Argiope t
20	27	45.0	60	2	AAV48316	Aav48316 Nucleotid
21	27	45.0	78	2	AAV48326	Aav48326 Nucleotid

C	22	27	45.0	156	10	ADK68635	Adk68635 HSP65-MUC
C	23	27	45.0	309	1	AAN90579	Aan90579 pDf9.3 CD
C	24	27	45.0	1194	12	AD157712	Ad157712 Human bre
C	25	27	45.0	1378	12	AD157693	Ad157693 Human bre
C	26	27	45.0	1424	12	ADO23180	Ado23180 Antisense
C	27	27	45.0	1428	6	ABL60159	Ab160159 Human MUC
C	28	27	45.0	1428	12	ADO23125	Ado23125 Human MUC
C	29	27	45.0	1457	12	ADF32627	Adf32627 Plasmid J
C	30	27	45.0	1527	2	AAV48329	Aav48329 MiniMUC1
C	31	27	45.0	1572	5	AAS00585	Aas00585 Human MUC
C	32	27	45.0	1614	12	ADK70370	Adk70370 Respirato
C	33	27	45.0	1630	12	AD157708	Ad157708 Human bre
C	34	27	45.0	1634	12	AD157689	Ad157689 Human bre
C	35	27	45.0	1712	12	AD157686	Ad157686 Human bre
C	36	27	45.0	1738	12	AD157669	Ad157669 Human bre
C	37	27	45.0	1755	12	AD157673	Ad157673 Human bre
C	38	27	45.0	1774	12	ADE43991	Ade43991 Plasmid J
C	39	27	45.0	1774	12	ADF32625	Adf32625 Plasmid J
C	40	27	45.0	1799	12	ADO23124	Ado23124 Human MUC
C	41	27	45.0	1800	10	ADK68629	Adk68629 HSP65-MUC
C	42	27	45.0	1803	12	AD157699	Ad157699 Human bre
C	43	27	45.0	1804	6	ABL67539	Ab167539 Thyroid c
C	44	27	45.0	1804	9	AAD56950	Aad56950 Human muc
C	45	27	45.0	1804	10	ADD14719	Add14719 Human src

ALIGNMENTS

RESULT 1
ID AAV48317 standard; DNA; 60 BP.
XX AAV48317;
AC 20-NOV-1998 (first entry)
DT
XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit R2.
XX
KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW tumour; tumour-associated antigen.
XX
OS Homo sapiens.
XX
XX PN WO9837095-A2.
XX PD 27-AUG-1998.
XX PF 24-FEB-1998; 98WO-US003693.
XX PR 24-FEB-1997; 97US-0038253P.
XX
PA (THER-) THERION BIOLOGICS CORP.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (DAND) DANA FARBER CANCER INST INC.
PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX WPI; 1998-467492/40.
XX
XX PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
XX immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
XX PS Disclosure; Page 11; 42pp; English.
XX
XX CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
XX immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
XX (RPV). The RPV was used in a pharmaceutical composition also containing
XX an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
XX recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
XX does not undergo significant genetic deletion, thereby providing an
XX unexpected stable and immunogenic pox virus. They can be used to
XX prevent or treat tumours expressing MUC1 tumour-associated antigens

```
XX SQ Sequence 60 BP; 17 A; 21 C; 13 G; 9 T; 0 U; 0 Other;
Query Match 100.0%; Score 60; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.6e-11;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCAGTACTGCACCGGACATGGCGTAACTACAGACCTGATACAAGACCTGCACCT 60
    |||||
Db 1 GGCAGTACTGCACCGGACATGGCGTAACTACAGACCTGATACAAGACCTGCACCT 60
    |||||

RESULT 2
ID AAV48325 standard; DNA, 60 BP.
XX AAV48325;
AC AAV48325;
DT 20-NOV-1998 (first entry)
DE Nucleotide sequence encoding MUC1 tandem repeat unit R10.
XX
KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KM tumour; tumour-associated antigen.
XX
OS Homo sapiens.
XX
PN WO9837095-A2.
XX
PD 27-AUG-1998.
XX
PF 24-FEB-1998; 98WO-US003693.
XX
PR 24-FEB-1997; 97US-0038253P.
XX
PA (THER-) THERION BIOLOGICS CORP.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA (DAND ) DANA FARBER CANCER INST INC.
XX
PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX
DR WPI; 1998-467492/40.
XX
PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS Disclosure; Page 11; 42pp; English.
XX
CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ Sequence 60 BP; 12 A; 22 C; 19 G; 7 T; 0 U; 0 Other;
Query Match 62.7%; Score 37.6; DB 2; Length 60;
Best Local Similarity 76.7%; Pred. No. 0.0016;
Matches 46; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 1 GGCAGTACTGCACCGGACATGGCGTAACTACAGACCTGATACAAGACCTGCACCT 60
    |||||
Db 1 GGTAGTACAGCGCCACCGGACATGGCGTACAGCGCTCCGATACGAGACCGGCGCCT 60
    |||||

RESULT 3
ID AAD00385 standard; DNA, 525 BP.
XX AAD00385;
AC AAD00385;
```

```
XX DT 29-AUG-2000 (first entry)
XX
DE Human Mucin 1 (MUC-1) protein fragment encoding DNA #2.
XX
KW Human; Mucin 1; MUC-1; tumour; pMRS30 expression vector; anti-tumour;
KW therapy; immune response; cytostatic; vaccine; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..525
FT FT /*tag= a
FT FT /product= "MUC-1 protein fragment"
XX
PN WO200025827-A2.
XX
PD 11-MAY-2000.
XX
PF 18-OCT-1999; 99WO-EP007874.
XX
PR 30-OCT-1998; 98IT-MI002330.
XX
PA (MENA ) MENARINI RICERCHE SPA.
XX
PI Parente D, Di Massimo AM, De Santis R;
XX
DR WPI; 2000-365410/31.
DR P-PSDB; AAY71021.
XX
PT Composition containing one or more DNA molecules encoding fragments of a
PT Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
PT tumor therapy.
XX
PS Claim 16; Fig 2; 56pp; English.
XX
CC The present sequence is a DNA encoding a fragment of human Mucin 1 (MUC-
CC 1) antigenic protein which is overexpressed in tumour cells. The sequence
CC was obtained from BT20 tumour cells by reverse transcriptase-PCR and
CC corresponds to nucleotides 205-720 of the EMBL sequence J05581 with a
CC start codon and two stop codons. The present sequence is cloned into a
CC pMRS30 expression vector and used in pharmaceutical composition e.g.
CC vaccine for inducing an antigen-specific anti-tumour immune response.
CC Composition containing this DNA molecule is useful in anti-tumour therapy
CC of patients affected with tumours characterised by high MUC-1 expression
XX
SQ Sequence 525 BP; 111 A; 199 C; 128 G; 87 T; 0 U; 0 Other;
Query Match 61.0%; Score 36.6; DB 3; Length 525;
Best Local Similarity 76.3%; Pred. No. 0.0061;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 1 GGCAGTACTGCACCGGACATGGCGTAACTACAGACCTGATACAAGACCTGCACCC 59
    |||||
Db 256 GGAAGTACTGCTCCACGACAGACAGCGGTGTACTCTCGGCTCCGATACCAAGCGCGGCC 314
    |||||

RESULT 4
ID AAD00391 standard; DNA, 891 BP.
XX AAD00391;
AC AAD00391;
DT 15-SEP-2003 (revised)
DT 29-AUG-2000 (first entry)
XX
DE Ubiquitin-E. coli Laci-human Mucin 1 fusion protein encoding DNA #2.
XX
KW Ubiquitin; Laci; beta-galactosidase; fusion protein; human; Mucin 1;
KW MUC-1; tumour; pMRS30 expression vector; anti-tumour; therapy;
KW immune response; cytostatic; vaccine; ds.
XX
OS Homo sapiens.
```

```
OS Escherichia coli.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 1..891
FT /tag= a
FT /product= "Ubiquitin-LacI-MUC-1 fusion protein"
FT misc_feature 1..369
FT /tag= b
FT /label= UBILACI DNA
FT /note= "includes ubiquitin-E. coli LacI fusion DNA"
FT misc_feature 370..891
FT /tag= c
FT /note= "Human MUC-1 partial DNA that corresponds to
FT nucleotides 205-720 of the EMBL sequence J05581 with two
FT stop codons"
XX
XX WO200025827-A2.
XX
XX 11-MAY-2000.
XX
XX 18-OCT-1999; 99WO-EP007874.
XX
XX 30-OCT-1998; 98IT-MI002330.
XX
XX (MENA ) MENARINI RICERCHE SPA.
XX
XX Parente D, Di Massimo AM, De Santis R;
XX WPI; 2000-365410/31.
XX DR P-PSDB; AAY71027.
XX
XX Composition containing one or more DNA molecules encoding fragments of a
XX Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
XX tumor therapy.
XX
XX Claim 18; Fig 8; 56pp; English.
XX
XX The present sequence is a DNA encoding a fusion protein consisting of
XX human Mucin 1 (MUC-1) fragment fused to UBILACI sequence at the N-
XX terminus. The UBILACI sequence consists of ubiquitin from MCF7 cell line
XX and a portion of E. coli beta-galactosidase (lacI). MUC-1 is an antigenic
XX protein overexpressed in tumour cells. The present sequence is cloned
XX into a pMRS30 expression vector and used in pharmaceutical composition
XX e.g. vaccine for inducing an antigen-specific anti-tumour immune
XX response. Composition containing this DNA molecule is useful in anti-
XX tumour therapy of patients affected with tumours characterised by high
XX MUC-1 expression. (Updated on 15-SEP-2003 to standardise OS field)
XX
XX Sequence 891 BP; 203 A; 307 C; 225 G; 156 T; 0 U; 0 Other;
SQ
Query Match 61.0%; Score 36.6; DB 3; Length 891;
Best Local Similarity 76.3%; Pred.No. 0.0069;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 1 GGCACTACTGCACCAACGGGACATGGCGTAAATCATCAGCACTGATACAAGACCTGCACC 59
DB 622 GGAAGTACTGCTCCACGACGACGATGTACCTCGGCTCCGATACCAAGCCGCCCC 680
RESULT 5
AAV48323
ID AAV48323 standard; DNA; 60 BP.
XX
XX AAV48323;
XX
XX 20-NOV-1998 (first entry)
XX
XX Nucleotide sequence encoding MUC1 tandem repeat unit R8.
XX
XX 58; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
XX tumour; tumour-associated antigen.
XX
```

```
OS Homo sapiens.
XX
XX WO9837095-A2.
XX
XX 27-AUG-1998.
XX
XX 24-FEB-1998; 98WO-US003693.
XX
XX 24-FEB-1997; 97US-0038253P.
XX
XX (THER-) THERION BIOLOGICS CORP.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX (DAND ) DANA FARBER CANCER INST INC.
XX
XX Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX WPI; 1998-467492/40.
XX
XX New recombinant pox virus for tumour therapy - comprises DNA encoding an
XX immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
XX Disclosure; Page 11; 42pp; English.
XX
XX The MUC1 tandem repeat units AAV48317-V48325 were used to create an
XX immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
XX (RPV). The RPV was used in a pharmaceutical composition also containing
XX an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
XX recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
XX does not undergo significant genetic deletion, thereby providing an
XX unexpectedly stable and immunogenic pox virus. They can be used to
XX prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
XX Sequence 60 BP; 10 A; 24 C; 17 G; 9 T; 0 U; 0 Other;
SQ
Query Match 59.0%; Score 35.4; DB 2; Length 60;
Best Local Similarity 79.2%; Pred.No. 0.0095;
Matches 42; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 7 ACTGCACCAACGGGACATGGCGTAAATCATCAGCACTGATACAAGACCTGCACC 59
DB 7 ACTGCCCTCCGGCGCATGTGTACCTCAGCTCGACACCAAGCCAGCCCC 59
RESULT 6
AAD00388
ID AAD00388 standard; DNA; 1371 BP.
XX
XX AAD00388;
XX
XX 29-AUG-2000 (first entry)
XX
XX Human Mucin 1 (MUC-1) protein fragment encoding DNA #5.
XX
XX Human; Mucin 1; MUC-1; tumour; pMRS30 expression vector; anti-tumour;
XX therapy; immune response; cytostatic; vaccine; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT CDS 1..1371
XX FT /tag= a
XX FT /product= "MUC-1 protein fragment"
XX
XX WO200025827-A2.
XX
XX 11-MAY-2000.
XX
XX 18-OCT-1999; 99WO-EP007874.
XX
XX 30-OCT-1998; 98IT-MI002330.
XX
XX (MENA ) MENARINI RICERCHE SPA.
XX
```

PI Parente D, Di Massimo AM, De Santis R;
XX
DR WPI; 2000-365410/31.
DR P-PSDB; AAY71024.
XX
PT Composition containing one or more DNA molecules encoding fragments of a
PT Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
PT tumor therapy.
XX
PS Claim 16; Fig 5; 56pp; English.
XX
CC The present sequence is a DNA encoding a fragment of human Mucin 1 (MUC-
CC 1) antigenic protein which is overexpressed in tumour cells. The sequence
CC was obtained by PCR from plasmids pMRS166, pMRS167, pMRS168 and pMRS169
CC which contain MUC-1 DNA from Br20 tumour cells. It corresponds to
CC nucleotides 136-1497 of the EMBL sequence J05581 with a start codon and
CC two stop codons. The present sequence is cloned into a pMRS30 expression
CC vector and used in pharmaceutical composition e.g. vaccine for inducing
CC an antigen-specific anti-tumour immune response. Composition containing
CC this DNA molecule is useful in anti-tumour therapy of patients affected
CC with tumours characterised by high MUC-1 expression
XX
SQ Sequence 1371 BP; 300 A; 450 C; 321 G; 300 T; 0 U; 0 Other;

Query Match 58.3%; Score 35; DB 3; Length 1371;
Best Local Similarity 74.6%; Pred. No. 0.027;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGCAGTACTGCACCAACCGGCATGGCGTTACATCAGACCTGATACAGACCTGCACC 59
Db 325 GGAAGTACCGCTCCACCAAGACACGCGTTTACCTCGGCTCCGATACCAAGCCGCCCC 383

RESULT 7
AAD00394
ID AAD00394 standard; DNA; 1737 BP.
XX
AC AAD00394;
XX
DT 15-SEP-2003 (revised)
DT 29-AUG-2000 (first entry)
XX
DE Ubiquitin-E. coli Laci-human Mucin 1 fusion protein encoding DNA #5.
XX
KW Ubiquitin; Laci; beta-galactosidase; fusion protein; human; Mucin 1;
KW MUC-1; tumour; pMRS30 expression vector; anti-tumour; therapy;
KW immune response; cytostatic; vaccine; ds.
XX
OS Homo sapiens.
OS Escherichia coli.
OS Chimeric.
XX
FH Key
FH CDS
FT
FT Location/Qualifiers
FT 1..1737
FT /tag= a
FT /product= "Ubiquitin-Laci-MUC-1 fusion protein"
FT 1..369
FT /tag= b
FT /label= UBILaci DNA
FT /note= "Includes ubiquitin-E. coli Laci fusion DNA"
FT 370..1737
FT /tag= c
FT /note= "Human MUC-1 partial DNA that corresponds to
FT nucleotides 136-1497 of the EMBL sequence J05581 with two
FT stop codons"
XX
XX
XX WO200025827-A2.
XX
XX 11-MAY-2000.
XX
XX
XX 18-OCT-1999; 99WO-EP007874.
XX
XX 30-OCT-1998; 98IT-MI002330.

XX
PA (MENA) MENARINI RICERCHE SPA.
XX
PI Parente D, Di Massimo AM, De Santis R;
XX
DR WPI; 2000-365410/31.
DR P-PSDB; AAY71030.
XX
PT Composition containing one or more DNA molecules encoding fragments of a
PT Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
PT tumor therapy.
XX
PS Claim 18; Fig 11; 56pp; English.
XX
CC The present sequence is a DNA encoding a fusion protein consisting of
CC human Mucin 1 (MUC-1) fragment fused to UBILaci sequence at the N-
CC terminus. The UBILaci sequence consists of ubiquitin from MCF7 cell line
CC and a portion of E. coli beta-galactosidase (Laci). MUC-1 is an antigenic
CC protein overexpressed in tumour cells. The present sequence is cloned
CC into a pMRS30 expression vector and used in pharmaceutical composition
CC e.g. vaccine for inducing an antigen-specific anti-tumour immune
CC response. Composition containing this DNA molecule is useful in anti-
CC tumour therapy of patients affected with tumours characterised by high
CC MUC-1 expression. (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ Sequence 1737 BP; 392 A; 558 C; 418 G; 369 T; 0 U; 0 Other;

Query Match 58.3%; Score 35; DB 3; Length 1737;
Best Local Similarity 74.6%; Pred. No. 0.029;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGCAGTACTGCACCAACCGGCATGGCGTTACATCAGACCTGATACAGACCTGCACC 59
Db 691 GGAAGTACCGCTCCACCAAGACACGCGTTTACCTCGGCTCCGATACCAAGCCGCCCC 749

RESULT 8
AAV48321
ID AAV48321 standard; DNA; 60 BP.
XX
AC AAV48321;
XX
DT 20-NOV-1998 (first entry)
DT
XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit R6.
XX
KW ss; MUC1; recombinant pox virus; cytotoxic T-Lymphocyte; immunogen;
KW tumour; tumour-associated antigen.
XX
OS Homo sapiens.
XX
PN WO9837095-A2.
XX
PD 27-AUG-1998.
XX
PF 24-FEB-1998; 98WO-US003693.
XX
PR 24-FEB-1997; 97US-0038253P.
XX
PA (THER-) THERION BIOLOGICS CORP.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (DAND) DANA FARBER CANCER INST INC.
XX
PI Schlom J, Kantar J, Kufe D, Panicali D, Gritz L;
XX
DR WPI; 1998-467492/40.
XX
XX
XX New recombinant pox virus for tumour therapy - comprises DNA encoding an
XX immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
XX Disclosure; Page 11; 42pp; English.
XX
XX The MUC1 tandem repeat units AAV48317-V48325 were used to create an

CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens
XX

SO Sequence 60 BP; 13 A; 27 C; 17 G; 3 T; 0 U; 0 Other;

Query Match 55.7%; Score 33.4; DB 2; Length 60;
Best Local Similarity 72.9%; Pred. No. 0.048;
Matches 43; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 GGCAGTACTGCACCGGACATGGCGTAAATCATCAGCAGCTGATACAGACCTGCACC 59
Db 1 GGCAGCAGCCGACCGCCCGCACACGGGGTCAACAAGCGCCGACACACTGACCTGCCCC 59

RESULT 9

AAV48322 ID AAV48322 standard; DNA; 60 BP.

XX AC AAV48322;

XX DT 20-NOV-1998 (first entry)

XX DE Nucleotide sequence encoding MUC1 tandem repeat unit R7.

XX KM SS; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
XX tumour; tumour-associated antigen.

XX OS Homo sapiens.

XX PN WO9837095-A2.

XX PD 27-AUG-1998.

XX PF 24-FEB-1998; 98WO-US003693.

XX PR 24-FEB-1997; 97US-0038253P.

XX PA (THER-) THERION BIOLOGICS CORP.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PA (DAND) DANA FARBER CANCER INST INC.

XX PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;

XX DR WPI; 1998-467492/40.

XX PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
XX immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX PS Disclosure; Page 11; 42pp; English.

CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens
XX

SO Sequence 60 BP; 14 A; 24 C; 16 G; 6 T; 0 U; 0 Other;

Query Match 55.7%; Score 33.4; DB 2; Length 60;
Best Local Similarity 72.9%; Pred. No. 0.048;
Matches 43; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 GGCAGTACTGCACCGGACATGGCGTAAATCATCAGCAGCTGATACAGACCTGCACC 59
Db 1 GGCAGTACCGCTCCACTGACACACGGGGTCAACAAGCGCCGACACTGACCTGCCCC 59

RESULT 10
AAV48320 ID AAV48320 standard; DNA; 60 BP.

XX AC AAV48320;

XX DT 20-NOV-1998 (first entry)

XX DE Nucleotide sequence encoding MUC1 tandem repeat unit R5.

XX KM SS; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
XX tumour; tumour-associated antigen.

XX OS Homo sapiens.

XX PN WO9837095-A2.

XX PD 27-AUG-1998.

XX PF 24-FEB-1998; 98WO-US003693.

XX PR 24-FEB-1997; 97US-0038253P.

XX PA (THER-) THERION BIOLOGICS CORP.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PA (DAND) DANA FARBER CANCER INST INC.

XX PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;

XX DR WPI; 1998-467492/40.

XX PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
XX immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX PS Disclosure; Page 11; 42pp; English.

CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens
XX

SO Sequence 60 BP; 9 A; 27 C; 15 G; 9 T; 0 U; 0 Other;

Query Match 55.3%; Score 33.2; DB 2; Length 60;
Best Local Similarity 75.9%; Pred. No. 0.056;
Matches 41; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 7 ACTGACACCGGACATGGCGTAAATCATCAGCAGCTGATACAGACCTGCACCT 60
Db 7 ACGGCCCCCTGCTCAAGGTGAATCCGCCCGGATACCAAGCCGGCCCT 60

RESULT 11

AAV48324 ID AAV48324 standard; DNA; 60 BP.

XX AC AAV48324;

XX DT 20-NOV-1998 (first entry)

XX DE Nucleotide sequence encoding MUC1 tandem repeat unit R9.

XX KM SS; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
XX tumour; tumour-associated antigen.

XX OS Homo sapiens.

This Page Blank (uspto)

Query Match 100.0%; Score 60; DB 6; Length 2297;
Best Local Similarity 100.0%; Pred. No. 4.3e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACTACTGCACCAACCGGCACATGGCGTAAATCATGACCACTGATACAAGACCTGCACCT 60
|||||
DB 406 GGCACTACTGCACCAACCGGCACATGGCGTAAATCATGACCACTGATACAAGACCTGCACCT 465
|||||

RESULT 2
BD225141
LOCUS BD225141 525 bp DNA linear PAT 17-JUL-2003
DEFINITION Medicinal composition having antitumor effect and containing DNA
encoding antigenic protein.
ACCESSION BD225141
VERSION BD225141.1 GI:33034911
KEYWORDS JP 2002528519-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 525)
AUTHORS Pallente,D., Massimo,A.M.D. and Desantis,R.
TITLE Medicinal composition having antitumor effect and containing DNA
JOURNAL Patent: JP 2002528519-A 2 03-SEP-2002;
MENARINI RICERCHE SPA
COMMENT OS Homo sapiens (human)
PN JP 2002528519-A/2
PD 03-SEP-2002
PF 18-OCT-1999 JP 2000579265
PR 30-OCT-1998 IT MI98A002330
PI DINO PALLENTE,ANNA MARIA D MASSIMO,RITA DESANTIS PC
A61K38/00,A61K35/76,A61K39/00,A61K48/00,A61P35/00,C12N15/09, PC
A61K37/02,
PC C12N15/00
CC Medicinal composition having antitumor effect and containing
CC DNA encoding
CC antigenic protein
CC Key Location/Qualifiers
FH Key 1..525
FT source /organism='Homo sapiens (human)'.
FEATURES
source 1..525
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 61.0%; Score 36.6; DB 6; Length 525;
Best Local Similarity 76.3%; Pred. No. 0.24;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 GGCACTACTGCACCAACCGGCACATGGCGTAAATCATGACCACTGATACAAGACCTGCACCT 59
|||||
DB 256 GGAACTACTGCTCCACCAAGACACACGCGTGTACCTCGGCTCCGATACCAAGCGCGCCCC 314
|||||

RESULT 3
BD225147
LOCUS BD225147 891 bp DNA linear PAT 17-JUL-2003
DEFINITION Medicinal composition having antitumor effect and containing DNA
encoding antigenic protein.
ACCESSION BD225147
VERSION BD225147.1 GI:33034917
KEYWORDS JP 2002528519-A/8.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 891)
AUTHORS Pallente,D., Massimo,A.M.D. and Desantis,R.

TITLE Medicinal composition having antitumor effect and containing DNA
JOURNAL Patent: JP 2002528519-A 8 03-SEP-2002;
MENARINI RICERCHE SPA
COMMENT OS Homo sapiens (human)
PN JP 2002528519-A/8
PD 03-SEP-2002
PF 18-OCT-1999 JP 2000579265
PR 30-OCT-1998 IT MI98A002330
PI DINO PALLENTE,ANNA MARIA D MASSIMO,RITA DESANTIS PC
A61K38/00,A61K35/76,A61K39/00,A61K48/00,A61P35/00,C12N15/09, PC
A61K37/02,
PC C12N15/00
CC Medicinal composition having antitumor effect and containing
CC DNA encoding
CC antigenic protein
CC Key Location/Qualifiers
FH Key 1..891
FT source /organism='Homo sapiens (human)'.
FEATURES
source 1..891
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 61.0%; Score 36.6; DB 6; Length 891;
Best Local Similarity 76.3%; Pred. No. 0.23;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 GGCACTACTGCACCAACCGGCACATGGCGTAAATCATGACCACTGATACAAGACCTGCACCT 59
|||||
DB 622 GGAACTACTGCTCCACCAAGACACACGCGTGTACCTCGGCTCCGATACCAAGCGCGCCCC 680
|||||

RESULT 4
BD225144
LOCUS BD225144 1371 bp DNA linear PAT 17-JUL-2003
DEFINITION Medicinal composition having antitumor effect and containing DNA
encoding antigenic protein.
ACCESSION BD225144
VERSION BD225144.1 GI:33034914
KEYWORDS JP 2002528519-A/5.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1371)
AUTHORS Pallente,D., Massimo,A.M.D. and Desantis,R.
TITLE Medicinal composition having antitumor effect and containing DNA
JOURNAL Patent: JP 2002528519-A 5 03-SEP-2002;
MENARINI RICERCHE SPA
COMMENT OS Homo sapiens (human)
PN JP 2002528519-A/5
PD 03-SEP-2002
PF 18-OCT-1999 JP 2000579265
PR 30-OCT-1998 IT MI98A002330
PI DINO PALLENTE,ANNA MARIA D MASSIMO,RITA DESANTIS PC
A61K38/00,A61K35/76,A61K39/00,A61K48/00,A61P35/00,C12N15/09, PC
A61K37/02,
PC C12N15/00
CC Medicinal composition having antitumor effect and containing
CC DNA encoding
CC antigenic protein
CC Key Location/Qualifiers
FH Key 1..1371
FT source /organism='Homo sapiens (human)'.
FEATURES
source 1..1371
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

[illegible]

Query Match	50.0%;	Score 30;	DB 3;	Length 80009;
Best Local Similarity	72.2%;	Pred. No. 19;		
Matches 39;	Conservative 0;	Mismatches 15;	Indels 0;	Gaps 0;
QY	7	ACTGCACCACCCGGACATGGCGTACATCAGCACCCTGATACAAGACCTGACACT	60	
Db	65954	ACGCCAGCACCAGCACCCTGACACCGACACCGACACTTGAGCCAACACCGACACT	65901	
RESULT 7				
AX180276/c				
LOCUS	AX180276	132 bp	DNA	linear
DEFINITION	Sequence 34 from Patent WO0146228.			PAT 06-AUG-2001
ACCESSION	AX180276			
VERSION	AX180276.1	GI:15132247		
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS	Crowe, J.S. and Ellis, J.H.			
TITLE	Nucleic acid vaccination			
JOURNAL	Patent: WO 0146228-A 34 28-JUN-2001;			

[illegible]

RESULT 8			
AX180275			
LOCUS	AX180275	144 bp	DNA
DEFINITION	Sequence 33 from Patent WO0146228.		linear
ACCESSION	AX180275		
VERSION	AX180275.1	GI:15132246	
KEYWORDS	.		
SOURCE	synthetic construct		
ORGANISM	synthetic construct		
	artificial sequences.		
REFERENCE	1		
AUTHORS	Crowe, J.S. and Ellis, J.H.		
TITLE	Nucleic acid vaccination		
JOURNAL	Patent: WO 0146228-A 33 28-JUN-2001;		
	GLAXO GROUP LIMITED (GB)		
FEATURES	Location/Qualifiers		
source	1..144		

```

/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Oligonucleotide"

```

```

Query Match          48.7%; Score 29.2; DB 6; Length 144;
Best Local Similarity 74.0%; Pred. No. 75;
Matches 37; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY      10 GCACCAACGGGACATGGCGTAATCATCAGACACCTGATACAAGACCTCACC 59
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db       37 GCTCACTCTGCACACGGCGGTGTTGCAGACACCTGCACACCGTCCGGCGCC 86

```

RESULT 9	AB017362/c	LOCUS	AB017362	557 bp	DNA	linear	INV 29-MAR-2000
DEFINITION	Bombyx mori gene for fibroin H-chain, partial cds.						
ACCESSION	AB017362						
VERSION	AB017362.1	GI:3599677					
KEYWORDS	fibroin H-chain; fibH.						
SOURCE	Bombyx mori (domestic silkworm)						
ORGANISM	Bombyx mori						
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea; Bombycidae; Bombyx.						
AUTHORS	1 (sites)						
TITLE	Tanaka, K., Kajiyama, N., Ishikura, K., Waga, S., Kikuchi, A., Ohtomo, K., Takagi, T. and Mizuno, S.						
JOURNAL	Determination of the site of disulfide linkage between heavy and light chains of silk fibroin produced by Bombyx mori						
MEDLINE	Biochim. Biophys. Acta 1432 (1), 92-103 (1999)						
PUBMED	99296390						
REFERENCE	10366732						
AUTHORS	2 (bases 1 to 557)						
	Tanaka, K., Kajiyama, N., Ishikura, K., Waga, S., Kikuchi, A., Takagi, T. and Mizuno, S.						


```

/db_xref="GI:765323"
/translacion="AGTSSSGRPYVANGVSCEYAWMSSESDPFGTSGAGSAGAGA
GSAGAGYGAGVGAGYGAGYGAAGACAGYAGAGSCVASGAGAGASGAGAGSAGAGS
GAGAGSVGAGSAGAGYGGAGAGYGGAGAGYGGAGAGYGGAGAGYGGAGAGYGGAGAGY
GAGAVGYGAGAGSAAASGAGAGSAGAGSAGAGSAGAGSAGAGSAGAGSAGAGSAGAGS
GSAGAGSAGAGYGGAGAGYGGAGAGSAGAGSAGAGSAGAGSAGAGSAGAGSAGAGSAG
GAGSAGAGSAGAGSAGAGSAGAGSAGAGSAGAGSAGAGSAGAGSAGAGSAGAGSAGAGY
GYGAGAGYGVGAGAGSAGAGSAGAGSAGAGSAGAGSAGAGSAGAGSAGAGSAGAGSAG
GAGSAGAGYGGAGYGGAGYGGAGYGGAGYGGAGYGGAGYGGAGYGGAGYGGAGYGGAGY
GAGSAGAGSAGAGSAGAGSAGAGSAGAGSAGAGSAGAGSAGAGSAGAGSAGAGSAGAGS
PGPYVANGYSRRREGYEYAWSKSPFTGTSAASGAGAGSAGAGSAGAGSAGAGSAGAGS
GSAGAGSAGSAGAGRYGGAGSAAASSVSASSRSYDYSRRTVRXNCGIPRLQLVVK
FRALPCVNC"

ORIGIN

Query Match      48.7%; Score 29.2; DB 3; Length 2008;
Best Local Similarity 74.0%; Pred. No. 55;
Matches 37; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Oy    11 CACCACCCGCATGCCCTTAACATCAGCACCTGTATAAAGACTGCACCT 60
     ||||| | | | | | | | | | | | | | | | | | | | | | | | | | 
Db   1718 CAGCAccgcacctgaaccagcaccgacctgtaacgcaccgccgacct 1669


RESULT 11
AE006311/c LOCUS          10701 bp DNA linear BCT 02-JUN-2004
DEFINITION Lactococcus lactis subsp. lactis IL1403 section 73 of 218 of the complete genome.
ACCESSION AE006311 AE005176
VERSION AE006311.1 GI:12723690
KEYWORDS .
SOURCE lactic acid bacterium
ORGANISM Lactococcus lactis subsp. lactis IL1403
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.

REFERENCE 1 (bases 1 to 10701) Bolotin,A., Wincker,P., Mauger,S., Jallion,O., Malarme,K., Weissenbach,J., Ehrlich,S.D. and Sorokin,A. The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp. lactis IL1403 Genome Res. 11 (5), 731-753 (2001)

TITLE Journal Medline Pubmed
JOURNAL Genom. Res. 11 (5), 731-753 (2001)
MEDLINE 21235186
PUBMED 11337471

AUTHORS 2 (bases 1 to 10701) Bolotin,A., Wincker,P., Mauger,S., Jallion,O., Malarme,K., Weissenbach,J., Ehrlich,S.D. and Sorokin,A. Direct Submission Submitted (09-JAN-2001) INRA, Genetique Microbiennne, Domaine de Vilvert, Jouy en Josas 78352, France

FEATURES source location/Qualifiers
1..10701 /organism="Lactococcus lactis subsp. lactis IL1403"
/mol_type="genomic DNA"
/strain="IL1403"
/sub_species="lactis"
/db_xref="taxon:272623"
114..335 /gene="acdPA"
/note="synonym: L0183"
114..335 /gene="acca"
/note="EVIDENCE BY HOMOLOG Y BIO07.01 FATTY ACID AND PHOSPHOLIPID METABOLISM. General KEY COMPONENT IN DE NOVO FATTY ACID BIOSYNTHESIS"
/codon_start=1
/product="acyl carrier protein"
/protein_id="AAK04870.1"
/db_xref="GI:12723691"
/translacion="MAVFEXVDIIVDELGEKEBEVTLETFSELADSLDLFQIIIND IEDFDVEVDTTEADMKTVAIDLKYVENNK"
615..1541
```

gene /gene="fabD"
/note="synonym: L0184"
615. .1541
/gene="fabD"
/note="EVIDENCE BY HOMOLOGY BIO07.01 FATTY ACID AND
PHOSPHOLIPID METABOLISM. General. FATTY ACID BIOSYNTHESIS"
/codon_start=1
/transl_table=11
/product="malonyl CoA-acyl carrier protein transacylase
(EC 2.3.1.39)"
/protein_id="AAK04871.1"
/db_xref="GI:12723692"
/translation="MTKTAFLFSQGAQKLGWARDLYDQETVKA TPDEASQALGYDL
RALIDNDEKLNETKYTPALITTSVAILRLSENGIKPDLVAGLSGEYSALVASGI
IDFQEAVKLVAKRGQYMTAAPTSGKMVAVMNTDPALIEILCOKAAEFKGIVSPAN
YMTPAQIVIGVEVADVAVELLKAGVRKLIELKVSQPFHTAILKPASEKALELDK
IDRRPELPLISNTSAKVMENDEKGLLTRQVMEPVRFYESVETMQKGATRFIEVGP
GRVLSGFIKIKDKNAEIANVENIASFEALINQ"
1627. .2358
/gene="fabG1"
/note="synonym: L0185"
1627. .2358
/gene="fabG1"
/note="EVIDENCE BY HOMOLOGY BIO07.01 FATTY ACID AND
PHOSPHOLIPID METABOLISM. General. FIRST REDUCTION STEP IN
THE FATTY ACID BIOSYNTHESIS PATHWAY"
/codon_start=1
/transl_table=11
/product="3-oxoacyl-acyl carrier protein reductase (EC
1.1.1.100)"
/protein_id="AAK04872.1"
/db_xref="GI:12723693"
/translation="MEIKKNVFTVGTSTRGICAIALQPAKAGSNLIINGSAISEEL
LAEFYAYGVKAVGISGDISKSDAKOMVAEITLGSVDILVNNAGITRDGLSKMSE
EDFESVLKINLTGAFNMTQAVLKPMTRARSGAIINISSVVGIMGNAGQVYASKAGL
IGFTKSIAREVARNVRNVAAPGFIESDMTEVLSDKVDAMKQIIPMKRFGMPEEIA
TATQFLAEQEVMTGQVLTIDGVSVM"
2399. .3628
/gene="fabF"
/note="synonym: L0186"
2399. .3628
/gene="fabF"
/note="EVIDENCE BY HOMOLOGY BIO07.01 FATTY ACID AND
PHOSPHOLIPID METABOLISM. General. FATTY ACID BIOSYNTHESIS"
/codon_start=1
/transl_table=11
/product="3-oxoacyl-acyl carrier protein synthase II (EC
2.3.1.41)"
/protein_id="AAK04873.1"
/db_xref="GI:12723694"
/translation="MTNRVVISGYGVSAIGNSPEEFWNLIKAGKTGICPITHPDAEA
TGISVAAEVKEFPDXYFOKKDARMDTFSLYAVYAALDAMEMSGITEENTFDRIGC
IMASGIGGLEBSQKNSQALIVAKGPRKGVAPLVPLAIANMATGNVALRTGARGVSRAE
VTACAGANSIGSAFREIKGYADAMIGAGAAICELGIAGFANLTALTKETDPEKA
CIPFDKNRSGFVMGEGSGVILLESLEHAQARGANILAEIVGYGNTNDAPHMTTPSGVG
AENAIKLALBEAGASPADVDYVNAHGTSTHANEETESKAIHNVIGDKAYVSTKALHG
HALGATGAIEAVATVQALLNQYAPVNA GTTELDEGMTINVVLGEGKPADINLALSDPF
GFGHNAVLAFFKWTGE"
3635. .4102
/gene="accB"
/note="synonym: L0187"
3635. .4102
/gene="accB"
/note="EVIDENCE BY HOMOLOGY BIO07.01 FATTY ACID AND
PHOSPHOLIPID METABOLISM. General. FIRST STEP IN LONG-CHAIN
FATTY ACID SYNTHESIS"
/codon_start=1
/transl_table=11
/product="biotin carboxyl carrier protein of acetyl-CoA
carboxylase (EC 6.4.1.2)"
/protein_id="AAK04874.1"
/db_xref="GI:12723695"
/translation="MNISEVKDLMNQFNNGSTLREFSWKNADGELSPSKNEGQALVGAQ

gene /gene="fabZ2"
4229. .4663
/gene="fabZ2"
/note="synonym: L0188"
4229. .4663
/gene="fabZ2"
/note="EVIDENCE BY HOMOLOGY BIO07.01 FATTY ACID AND
PHOSPHOLIPID METABOLISM. General. FATTY ACID BIOSYNTHESIS"
/codon_start=1
/transl_table=11
/product="3R-hydroxymyristoyl-(3R)-HYDROXYMYRISTOYL-acyl
carrier protein dehydratase (EC 4.2.1.1)"
/protein_id="AAK04875.1"
/db_xref="GI:12723696"
/translation="MTEVINVTIMEALPHRYPELLVDRVIDIAEDEITAIKNTIN
BEFQGHFPQYPVMPGVLMELAQAGVLELSKRENKGLVFYAGMDNVKIKQVTP
GDQLVHAKEFIKRRGP IAVEAEATVDGLAKGTLTFALGR"
5057. .6424
/gene="accC"
/note="synonym: L0189"
5057. .6424
/gene="accC"
/note="EVIDENCE BY HOMOLOGY BIO07.01 FATTY ACID AND
PHOSPHOLIPID METABOLISM. General. FIRST STEP IN LONG-CHAIN
FATTY ACID SYNTHESIS"
/codon_start=1
/transl_table=11
/product="biotin carboxylase (EC 6.3.4.14)"
/protein_id="AAK04876.1"
/db_xref="GI:12723697"
/translation="MFNKILINRGEIARIRARELGVGTVAIYSTADRALHVL
ADEAICVGPARTSESYLNMNNTLISAATGAQAIHPGFISENSKFARLCENNIKF
IGBSAKVMDMCKINARAEMTKAGVPVTPGSEGEVHTTDEAVKIAEIGYIMLKAS
AGGGKGIRKVNVEELVPAPFAQAQAAFGNGAMFIERMIFPARHIEVQILGDSQ
GNVTHLGERDCSLORNNQKVLSESPVAIGHTLRLNEITAAVAKAAQHVSYENAGTIEF
LIDESSGKFYEMEMNTRVQVEHPTEFVSQVQIVKEQIRVANGNELSVKQEVTFESGH
AIECRINAENPKENFAPSPGKITNLFPSGGVGLRVDSAMYSGYSIPPIYSMIAKVI
VHGENRFEALMKQRALLEFVEGVITNVDFQLELISDPHVAAGDYDTSFLGNVFLPE
YQXED"
6510. .7376
/gene="accD"
/note="synonym: L0180"
6510. .7376
/gene="accD"
/note="EVIDENCE BY HOMOLOGY BIO07.01 FATTY ACID AND
PHOSPHOLIPID METABOLISM. General. FATTY ACID BIOSYNTHESIS"
/codon_start=1
/transl_table=11
/product="acetyl-CoA carboxylase carboxyl transferase
subunit beta (EC 6.4.1.2)"
/protein_id="AAK04877.1"
/db_xref="GI:12723698"
/translation="MALFOKKYIKIKINPNRSIIIEKQAEQEVVDELFAKCPACKHTIY
QKDLGNKVCNCDYNFRITAKERLAIIVADKDSFVEMFTGESKNPLDFPGYPEKLA
TKARTGLDEAVITGTATIKOKTALAIMDSTFIMASMGIVVEKLTRLFEYATTEKLP
IIVFTASGGARMOEGIMQAKTSAAYKRHSNAGLFTYITVLTDPPTGGVTASFASL
GDIILAEQSLIGFAGRVITEQTVRQTLPPDDFOKAEFLNLNHFVDAYIKRTELROKLA
LLELHTEVENY"
7369. .8151
/gene="accA"
/note="synonym: L0190"
7369. .8151
/gene="accA"
/note="EVIDENCE BY HOMOLOGY BIO07.01 FATTY ACID AND
PHOSPHOLIPID METABOLISM. General. FIRST STEP IN LONG-CHAIN
FATTY ACID SYNTHESIS"
/codon_start=1
/transl_table=11
/product="acetyl-CoA carboxylase carboxyl transferase
subunit alpha (EC 6.4.1.2)"
/protein_id="AAK04878.1"
/db_xref="GI:12723699"

gene
/translation="MSETIDSLTGEVSQKNIADIINEARADRIARDITDKVFTDF
ELHGRDLTDEDAIVGIGARFNNOAVTVIGIOKGRNLAEENLATNFGSPNGYRKALR
LMQAEKFGFRPVITFTINAGAYPGIEABERGCGEAIARNLLEMSDLKPIIATITGEG
GSGGALALANKVYILENAVSVLSPEGFATILMKDGRDRDEAAELMKITAPHLLEM
GLVDATISEENLRQLTESLGLDSLTTETELAEQRERYERFKY"
8280..9422
/gene="metB2"

Query Match 47.7%; Score 28.6; DB 1; Length 10701;
Best Local Similarity 72.5%; Pred. No. 70;
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 10 GCACCAACCGGCATGGCGTATACATCAGCACCTGATACAGACCTGCACCT 60
|||||
Db 2901 GCACCAACCGGCATGCGTACTTCAAGCAGCGAAACCCACGAGCACCT 2851

RESULT 12
AC119248 227395 bp DNA linear ROD 29-APR-2004
LOCUS Mus musculus chromosome 8, clone RP24-273H15, complete sequence.
DEFINITION AC119248
AC119248.7 GI:46849687
VERSION HTG.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Birren, B., Nussbaum, C. and Lander, E.
TITLE 1 (bases 1 to 227395)
JOURNAL Mus musculus chromosome 8, clone RP24-273H15
REFERENCE
AUTHORS Unpublished
TITLE 2 (bases 1 to 227395)

Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L.,
Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,
Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., MacDonald, P., Major, J., Margulis, N., Matthews, C.,
McCarthy, M., McGowan, P., McKernan, K., Meldrim, J., Meneus, L.,
Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (25-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 227395)

Birren, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, M., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T.,
Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K.,
Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C.,
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,

Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Strange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (18-MAR-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 227395)

REFERENCE
AUTHORS Birren, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, M., Anderson, S., Arachchi, H. M., Barna, N., Bastien, V.,
Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J.,
Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B.,
Dearellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L.,
Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D.,
Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N.,
Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I.,
Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T.,
Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R.,
Maclean, C., MacDonald, P., Major, J., Manning, J., Matthews, C.,
McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C.,
O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupack, R.,
Seaman, S., Severy, P., Smith, C., Spencer, B., Strange-Thomann, N.,
Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodore, J.,
Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R.,
Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.,
Zimmer, A. and Zody, M.
Direct Submission
Submitted (29-APR-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 29, 2004 this sequence version replaced gi:45544751.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

COMMENT
TITLE Direct Submission
JOURNAL
COMMENT Center: Whitehead Institute/MIT Center for Genome Research
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@broad.mit.edu
Project Information
Center project name: L25340
Center clone name: 273_H_15

FEATURES
source

Location/Qualifiers
1..227395
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="8"
/map="8"
/clone="RP24-273H15"
/clone_lib="RPCL-24 Male Mouse BAC"
1...6920
/note="wgs end extension
clone end:5P6"
8..91
/rpt_family="CT-rich"
complement(114..1557)
/rpt_family="RMR6B"
2654..2699
/rpt_family="(CTTG)n"
2710..3134
/rpt_family="WTC"
3135..5529
/rpt_family="L1P5"
3821..3842
/note="single clone coverage"
5834..6010
/rpt_family="B2_Mm2"

Page 8

```
repeat_region      6032. .6296      /rpt_family="L1M2"
repeat_region      6426. .6461      /rpt_family="(CA)n"
repeat_region      complement(6700. .6853)
                    /rpt_family="RLTR11A"
repeat_region      6903. .6979      /rpt_family="L1ME"
misc_feature        6921. .6924      /note="clone_boundary
                    clone end:SP6
                    site:MboI"
repeat_region      7330. .7378      /rpt_family="(TTTGTG)n"
repeat_region      complement(7381. .7935)
                    /rpt_family="Lx"
repeat_region      complement(8566. .8781)
                    /rpt_family="MLT1D"
repeat_region      complement(9071. .9197)
                    /rpt_family="ORR1D"
repeat_region      9384. .9442      /rpt_family="(TTAA)n"
repeat_region      9415. .9518      /rpt_family="(TATAA)n"
repeat_region      complement(9586. .9807)
                    /rpt_family="URR1B"
repeat_region      10143. .10171      /rpt_family="(A)n"
repeat_region      10201. .10335      /rpt_family="B1F"
repeat_region      11649. .11826      /rpt_family="MER20"
repeat_region      11897. .12045      /rpt_family="(TC)n"
repeat_region      12610. .12654      /rpt_family="GA-rich"
repeat_region      14438. .14460      /rpt_family="(TTTG)n"
repeat_region      15295. .15386      /rpt_family="(TAA)n"
repeat_region      15527. .15561      /rpt_family="(TTTGTG)n"
repeat_region      complement(15755. .15820)
                    /rpt_family="B3"
repeat_region      complement(16742. .16891)
                    /rpt_family="B1F"
repeat_region      complement(16827. .16905)
                    /rpt_family="B4A"
repeat_region      17239. .17269      /rpt_family="(TGG)n"
repeat_region      complement(18173. .18259)
                    /rpt_family="L3b"
repeat_region      18626. .20821      /rpt_family="L1_MM"
repeat_region      22232. .22273      /rpt_family="(TAA)n"
repeat_region      22855. .22915      /rpt_family="(GGGA)n"
repeat_region      complement(24184. .24315)
                    /rpt_family="ID_B1"
repeat_region      complement(24808. .25550)
                    /rpt_family="RMER1B"
repeat_region      27219. .27345      /rpt_family="B4A"
repeat_region      27346. .27366      /rpt_family="(CAGAGA)n"
repeat_region      27367. .27491      /rpt_family="B4A"
repeat_region      complement(29732. .29960)
                    /rpt_family="L1MB4"
repeat_region      complement(29941. .30530)
                    /rpt_family="MER82"
repeat_region      30805. .30919
```

Query Match	47.7%	Score 28.6	DB 10	Length 227395
Best Local Similarity	67.8%	Pred. No. 49		
Matches	40	Conservative	0	Mismatches 19, Indels 0, Gaps 0
QY	1	GGCAGTACTGCACCA	CGGCATCGGTACATCAGCACCCTGATACAGACCTGCACC	59
Db	157274	GCCAGCACTCCACCA	GAGCAGCAGTCATGATGCAATCAGCAGTGCTATAATACCTTGCCC	157332
RESULT 13				
AC109139/c		256754 bp	DNA	Linear HTG 13-MAR-2004
LOCUS		Mus musculus chromosome 8 clone RP23-47L13	map 8,	WORKING DRAFT
DEFINITION		SEQUENCE, 27 unordered pieces.		
AC109139		AC109139	GI:45430241	
VERSION		AC109139.4	GI:45430241	
KEYWORDS		HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.		
SOURCE		Mus musculus (house mouse)		
ORGANISM		Mus musculus		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE		1 (bases 1 to 256754)		
JOURNAL		Birren, B., Nusbaum, C. and Lander, E.		
REFERENCE		Mus musculus chromosome 8, clone RP23-47L13		
AUTHORS		Unpublished		
TITLE		2 (bases 1 to 256754)		
JOURNAL		Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,		
REFERENCE		Boeckmann, J., Bouckgatter, B.,		
AUTHORS		Boeckmann, J., Bouckgatter, B.,		

TITLE	REFERENCE
Direct Submission	Submitted (03-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 256754)	
Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgelter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, D., Gadya, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, R., Maclean, C., Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Meneses, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiljev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.	

TITLE
JOURNAL
COMMENT

Direct Submission
Submitted (13-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 13, 2004 this sequence version replaced gi:21326298.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
Center project name: L14110
Center clone name: 47_L13

----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 246913 bases at least Q40
Consensus quality: 250922 bases at least Q30
Consensus quality: 252476 bases at least Q20
Insert size: 275000; agarose-fp
Insert size: 253647; sum-of-contigs
Quality coverage: 5.9 in Q20 bases; agarose-fp
Quality coverage: 6.4 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 27 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1244: contig of 1244 bp in length
* 1245 1344: gap of 100 bp
* 1345 2423: contig of 1079 bp in length
* 2424 2523: gap of 100 bp
* 2524 18381: contig of 15858 bp in length
* 18382 18481: gap of 100 bp
* 18482 20692: contig of 2211 bp in length
* 20693 20792: gap of 100 bp
* 20793 22569: contig of 1777 bp in length
* 22570 22670: gap of 100 bp
* 22670 37683: contig of 15014 bp in length
* 37684 37783: gap of 100 bp
* 37784 62795: contig of 25012 bp in length
* 62796 62895: gap of 100 bp
* 62896 117535: contig of 54640 bp in length
* 117536 117635: gap of 100 bp
* 117636 119189: contig of 1554 bp in length
* 119190 119289: gap of 100 bp
* 119290 121556: contig of 2267 bp in length
* 121557 121656: gap of 100 bp
* 121657 122227: contig of 571 bp in length
* 122228 122327: gap of 100 bp
* 122328 123313: contig of 986 bp in length
* 123314 123413: gap of 100 bp
* 123414 128705: contig of 5292 bp in length
* 128706 128805: gap of 100 bp
* 128806 130201: contig of 1396 bp in length
* 130202 130301: gap of 100 bp
* 130302 134936: contig of 4635 bp in length
* 134937 135036: gap of 100 bp
* 135037 138419: contig of 3383 bp in length
* 138420 138519: gap of 100 bp
* 138520 145823: contig of 7304 bp in length
* 145824 145923: gap of 100 bp
* 145924 180246: contig of 34323 bp in length
* 180247 180346: gap of 100 bp
* 180347 183714: contig of 3368 bp in length
* 183715 183814: gap of 100 bp

* 183815 209464: contig of 25650 bp in length
* 209465 209564: gap of 100 bp
* 209565 214498: contig of 4934 bp in length
* 214499 214598: gap of 100 bp
* 214599 216080: contig of 1482 bp in length
* 216081 216180: gap of 100 bp
* 216181 249862: contig of 33682 bp in length
* 249863 249962: gap of 100 bp
* 249963 253576: contig of 3614 bp in length
* 253577 253676: gap of 100 bp
* 253677 254277: contig of 601 bp in length
* 254278 254377: gap of 100 bp
* 254378 256125: contig of 1748 bp in length
* 256126 256225: gap of 100 bp
* 256226 256754: contig of 529 bp in length.

FEATURES
Source
Location/Qualifiers
1. .256754
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="8"
/map="8"
/clone="RP23-47L13"
/clone_11b="RPCI-23 Female Mouse BAC"
1. .1244
/note="assembly_fragment"
1345. .2423
/note="assembly_fragment"
2524. .18381
/note="assembly_fragment"
18482. .20692
/note="assembly_fragment"
20793. .22569
/note="assembly_fragment"
22670. .37683
/note="assembly_fragment"
37784. .62795
/note="assembly_fragment"
62896. .117535
/note="assembly_fragment"
117636. .119189
/note="assembly_fragment"
119290. .121556
/note="assembly_fragment"
121657. .122227
/note="assembly_fragment"
122328. .123313
/note="assembly_fragment"
123414. .128705
/note="assembly_fragment"
128806. .130201
/note="assembly_fragment"
130302. .134936
/note="assembly_fragment"
135037. .138419
/note="assembly_fragment"
138520. .145823
/note="assembly_fragment"
145924. .180246
/note="assembly_fragment"
180347. .183714
/note="assembly_fragment"
183815. .209464
/note="assembly_fragment"
209565. .214498
/note="assembly_fragment"

Query Match 47.7%; Score 28.6; DB 2; Length 256754;
Best Local Similarity 67.8%; Pred. No. 48;
Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

OY 1 GGCAGTACTGACCAACCGACATGCGTAACATCAGACCTGATACAGACCTGCACC 59
| ||| ||| ||||| ||||| | | ||||| | || ||| |

DB	108445	GCCAGCACTCCACCAAGGACAGTCATGCATCAGACAGTGCTATATACTTGCCC	108387
RESULT 14			
AC109778			
LOCUS			
DEFINITION	AC109778	69333 bp	DNA linear PRI 31-MAY-2002
			Homo sapiens 3 BAC RP11-17013 (Roswell Park Cancer Institute Human
			BAC library) complete sequence.
ACCESSION	AC109778		
VERSION	AC109778.8	GI:21281345	
KEYWORDS	HTG.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 69333)		
AUTHORS	Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbara,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Checko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Dem,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landy,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., Mcleod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabat,K., Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I., Sodergren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatok,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstein,G. and Gibbs,R.		
TITLE	Direct Submission		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 69333)		
AUTHORS	Worley,K.C.		
TITLE	Direct Submission		
JOURNAL	Submitted (07-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
REFERENCE	3 (bases 1 to 69333)		
AUTHORS	Worley,K.C.		
TITLE	Direct Submission		
JOURNAL	Submitted (26-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
REFERENCE	4 (bases 1 to 69333)		
AUTHORS	Worley,K.C.		

TITLE	Direct Submission
JOURNAL	Submitted (31-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT	On May 31, 2002 this sequence version replaced gi:20514664. INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email gc-help@bcm.tmc.edu
CLONE LENGTH:	This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the features listing.
ANNOTATION OF FEATURES:	STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts. Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences. Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.
SEQUENCING READ COVERAGE:	Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as low Coverage.
QUALITY OF INDIVIDUAL BASES:	This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html .
FEATURES	QUALSTAT-REPORT.
source	Location/Qualifiers 1. .69333 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="3" /clone="RP11-17013" /note="complement(1. .2006) /function="overlaps bases 1. .2006 of clone AC069529"
misc_feature	
repeat_region	1844. .1864 /rpt_family="AT_rich" complement(2195. .2250) /rpt_family="MER102"
repeat_region	2862. .2867 /rpt_family=" (CA) n" 2868. .2900 /rpt_family=" (CA) n" 3085. .3192 /rpt_family=" (CA) n"
repeat_region	3193. .3504 /rpt_family="HAL1" 3505. .3771 /rpt_family="AluY"
repeat_region	4324. .4496 /rpt_family="HAL1" 5007. .5167 /rpt_family="MER5B" complement(5194. .5964) /rpt_family="MER99"
repeat_region	

Query Match	47.3%	Score 28.4	DB 9	Length 69333
Best Local Similarity	70.4%	Pred. No. 65		
Matches 38	Conservative 0	Mismatches 16	Indels 0	Gaps 0
Db	69086	GACATTCTCTGCACCACTGGCTCATGATGCACTTCAGCAGCATTAACAAACAT	69139	
LOCUS	AC117438/c	96545 bp	DNA	linear
DEFINITION	Homo sapiens 3.BAC RP11-190F5 (Roswell Park Cancer Institute Human			PRI 26-JUN-2002

ACCESSION VERSION KEYWORDS SOURCE ORGANISM	BAC Library) complete sequence.
AC117438 AC117438.4 HTG.	GI:21591816
Homo sapiens (human)	
Homo sapiens	
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE AUTHORS	Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Alsbrooks, S.L., Amarantunge, H.C., Are, J.R., Ayele, M., Banks, T., Barbaria, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flags, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., He, X., Hernandez, J., Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B., Homsl, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Liu, C., Liu, J., Liu, W., Loulesged, H., Lozada, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Moore, S., Morgan, M., Moorish, T., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenko, S., Oguh, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L., Quiles, M., Ren, Y., Rives, M., Scott, G., Shen, H., Shooshari, N., Sisson, I., Sodergren, E., Sonake, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczek, K., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Naylor, S.L., Weinstock, G. and Gibbs, R.
REFERENCE AUTHORS	Direct Submission
REFERENCE TITLE	Unpublished
REFERENCE JOURNAL	2 (bases 1 to 96545)
REFERENCE AUTHORS	Worley, K.C.
REFERENCE TITLE	Direct Submission
REFERENCE JOURNAL	Submitted (10-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE AUTHORS	3 (bases 1 to 96545)
REFERENCE TITLE	Worley, K.C.
REFERENCE JOURNAL	Direct Submission
REFERENCE AUTHORS	Submitted (23-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE TITLE	4 (bases 1 to 96545)
REFERENCE JOURNAL	Worley, K.C.
REFERENCE AUTHORS	Direct Submission
REFERENCE TITLE	Submitted (26-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE JOURNAL	On Jun 26, 2002 this sequence version replaced gi:21490104.
REFERENCE AUTHORS	INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STRs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases.

Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found

at URL:

http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

FEATURES

source

1. 96545
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-190F5"
1380. 1401
/rpt_family="(T)n"
complement(2723. 3026)
/rpt_family="AluJo"
3376. 3629
/rpt_family="MIR"
5318. 5564
/rpt_family="L2"
5754. 6117
/rpt_family="MLT1B"
8811. 9638
/rpt_family="L2"
10070. 10296
/rpt_family="MER3"
complement(11148. 11297)
/rpt_family="MER5A"
12694. 13521
/rpt_family="L2"
13697. 13811
/rpt_family="L2"
complement(13957. 14361)
/rpt_family="MLT1D"
14414. 14469
/rpt_family="L2"
14850. 15213
/rpt_family="MLT1E"
complement(15861. 15933)
/rpt_family="MIR"
complement(16333. 16368)
/rpt_family="MIR"
16648. 16706

/rpt_family="MIR"
16823. 16845
/rpt_family="(TC)n"
16845. 16899
/rpt_family="(TA)n"
complement(17142. 17251)
/rpt_family="L2"
complement(17832. 17893)
/rpt_family="L1PA12"
17894. 18012
/rpt_family="L1PA13"
18420. 18702
/rpt_family="AluSp"
18703. 18730
/rpt_family="(CAA)n"
19326. 19355
/rpt_family="AT_rich"
19364. 19413
/rpt_family="MIR"
20916. 21157
/rpt_family="AluSg"
complement(21560. 21933)
/rpt_family="MSTC"
complement(24710. 25383)
/rpt_family="L2"
complement(25880. 26180)
/rpt_family="AluYb8"
complement(26723. 26814)
/rpt_family="MER81"
complement(27544. 27741)
/rpt_family="MER8"
28231. 28707
/rpt_family="MLT1D"
28957. 29110
/rpt_family="L1ME"
29557. 29728
/rpt_family="AluJo/FRAM"
29780. 29814
/rpt_family="(TTCA)n"
29969. 30092
/rpt_family="MIR"
31848. 31876
/rpt_family="AT_rich"
32282. 32318
/rpt_family="(TG)n"
33871. 33910
/rpt_family="(TTTAA)n"
complement(33913. 34060)
/rpt_family="L1MA4"
34065. 34144
/rpt_family="L1MA4"
34145. 34345

Query Match 47.3%; Score 28.4; DB 9; Length 96545;
Best Local Similarity 70.4%; Pred.No.63;
Matches 38; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 GGCAGTACTGCACCAACCGCACATGGCGTAAACATCAGCACTGTATCAAGACCT 54
DB 94789 GACATTCTCGACCAACTGGCTCATGATGCAACTTCAGCAGCATAAACAAACAT 94736

Search completed: January 15, 2005, 22:36:45
Job time : 725.8 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 17:40:52 ; Search time 722.8 Seconds
(without alignments)
3925.542 Million cell updates/sec

Title: US-10-057-136-2

Perfect score: 60
Sequence: 1 ggctccaccgccccccagc.....cggacaccagcgccgccccg 60

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_stb:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	60	100.0	120	6 AX192396	AX192396 Sequence
C 2	60	100.0	120	6 BD000571	BD000571 Human pol
3	60	100.0	1414	12 AF423031	AF423031 Synthetic
4	60	100.0	1455	6 CQ715242	CQ715242 Sequence
5	60	100.0	1457	6 AX959914	AX959914 Sequence
6	60	100.0	1774	6 AX959912	AX959912 Sequence
7	60	100.0	1804	6 AR492318	AR492318 Sequence
8	60	100.0	1804	6 AX335367	AX335367 Sequence
9	60	100.0	1804	6 HUMMUCAB	J05581 Human polym
10	60	100.0	1834	12 AF423030	AF423030 Synthetic
11	60	100.0	1835	6 AX959918	AX959918 Sequence
12	60	100.0	2135	6 AX959916	AX959916 Sequence
13	60	100.0	2297	6 BD272907	BD272907 A recombi
14	60	100.0	4139	6 CQ834017	CQ834017 Sequence
15	60	100.0	4139	6 AX334899	AX334899 Sequence
16	60	100.0	4139	6 AX335372	AX335372 Sequence
17	60	100.0	4139	6 AX336712	AX336712 Sequence
18	60	100.0	4139	6 AX409474	AX409474 Sequence
19	60	100.0	4139	6 AX409474	AX409474 Sequence

20	60	100.0	4139	6 AX440481	AX440481 Sequence
21	60	100.0	4139	9 HUMPAWNU	J05582 Human pancr
22	60	100.0	7188	9 AY463543	AY463543 Homo sapi
23	60	100.0	8181	6 AX406624	AX406624 Sequence
24	60	100.0	8181	9 HUMPEM	M61170 Human polym
25	60	100.0	8186	6 AR492320	AR492320 Sequence
C 26	60	100.0	133525	9 AL713999	AL713999 Human DNA
27	58.4	97.3	120	6 AX192397	AX192397 Sequence
28	58.4	97.3	120	6 BD000572	BD000572 Human pol
29	58.4	97.3	572	6 AR492319	AR492319 Sequence
30	58.4	97.3	572	9 HUMDF3AA	M31823 Human breas
31	58.4	97.3	1721	6 CQ771290	CQ771290 Sequence
32	58.4	97.3	1721	6 AR492306	AR492306 Sequence
33	58.4	97.3	1721	6 AX335860	AX335860 Sequence
34	58.4	97.3	1721	6 AX440427	AX440427 Sequence
35	58.4	97.3	1721	6 AX587588	AX587588 Sequence
36	58.4	97.3	1721	9 HSTBYMA	X52229 Human mRNA
37	58.4	97.3	1971	6 AX963157	AX963157 Sequence
38	58.4	97.3	2037	6 AX963159	AX963159 Sequence
39	58.4	97.3	2238	9 HSSETA	X52228 Human mRNA
40	57	95.0	491	9 HUMEPISIB1	M32739 Human episi
41	57	95.0	518	6 AR492402	AR492402 Sequence
42	57	95.0	518	9 HUMEPISIA1	M32738 Human episi
43	56.8	94.7	180	6 E08763	E08763 CDNA encodi
44	56.8	94.7	1572	6 AX093798	AX093798 Sequence
45	55.4	92.3	1320	6 E08764	E08764 CDNA encodi

ALIGNMENTS

RESULT 1	AX192396/c	120 bp	DNA	linear	PAT 15-AUG-2001
LOCUS	AX192396	Sequence 2 from Patent EPI103623.			
DEFINITION	AX192396				
ACCESSION	AX192396.1	GI:15210363			
VERSION					
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 Taylor-Papadimitriou,J., Burchell,J. and Gendler,S.				
AUTHORS	Human mucin core protein: nucleic acid probes, peptide fragments				
TITLE	and antibodies thereto, and uses thereof in diagnostic and				
JOURNAL	therapeutic methods				
FEATURES	Patent: EP 1103623-A 2 30-MAY-2001;				
source	IMPERIAL CANCER RESEARCH TECHNOLOGY LIMITED (GB)				
ORIGIN	Location/Qualifiers				
	1..120				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
QY	1 GGCTCCACCGCGCCCGCCAGCCGCGTGTACCTCGCGCCCGGACACCGCGCGCCCG 60				
Db	87 GGCTCCACCGCGCCCGCCAGCCGCGTGTACCTCGCGCCCGGACACCGCGCGCCCG 28				
RESULT 2	BD000571/c	120 bp	DNA	linear	PAT 31-JAN-2002
LOCUS	BD000571	Human polymorphic epithelial mucin core protein and nucleic acid			
DEFINITION	BD000571	encoding the protein.			
ACCESSION	BD000571.1	GI:18623684			
VERSION					
KEYWORDS	JP 2000333675-A/2.				

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 120)
TITLE Papadimitrov,J.T., Jendora,S. and Bachieru,J.
JOURNAL Human polymorphic epithelial mucin core protein and nucleic acid encoding the protein
COMMENT Patent: JP 2000333675-A 2 05-DEC-2000; IMPERIAL CANCER RESEARCH TECHNOLOGY LTD
OS Homo sapiens (human)
PN JP 2000333675-A/2
PD 05-DEC-2000
PR 26-APR-2000 JP 2000125724
PR 07-JAN-1987 GB 8700269,07-JAN-1987 GB 8700279 PR
22-APR-1987 US 041306,09-NOV-1987 GB 8726172 PI JOYCE
TAYLOR PAPADIMITROV,SANDRA JENDORA,JOY BACHIERU PC
C12N15/02,A61K38/00,A61K39/395,A61K49/00,A61P35/00, PC
C07K14/47,
PC C07K16/44,C12N5/10,C12P21/08//C12N15/02,C12R1:91),(C12N5/10,
PC C12R1:91),
PC C12N15/00,A61K37/02,C12N5/00,(C12N15/00,C12R1:91),(C12N5/00,
PC C12R1:91)
CC
FH Key Location/Qualifiers
FT source 1..120
FEATURES Location/Qualifiers
source 1..120
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 60; DB 6; Length 120;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTCCACCGCCCCCAGCCCGACGGTGTCACTCGGCCCCGACACCGCGCCCCCG 60
Db 87 GGCTCCACCGCCCCCAGCCCGACGGTGTCACTCGGCCCCGACACCGCGCCCCCG 28
RESULT 3 1414 bp mRNA linear SYN 10-JUL-2003
AF423031
LOCUS Synthetic construct Homo sapiens mucin variant MUC1-CT58 (MUC1)
DEFINITION mRNA, complete cds; alternatively spliced.
ACCESSION AF423031
VERSION AF423031.1 GI:19338621
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..1414
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"

/note="Homo sapiens gene in transgenic Mus musculus C57BL/6; isolated from intestinal mucosa"
1..1414
/gene="MUC1"
1..1386
/gene="MUC1"
/note="alternatively spliced; contains exon 6b resulting in variant carboxy-terminal domain; lacks sites for beta-catenin and Grb2 interactions; derived from Homo sapiens"
/codon_start=1
/transl_table=11
/product="mucin variant MUC1-CT58"
/protein_id="AA186735.1"
/db_xref="GI:19338622"
/translation="MTPEGTQSPFFLLLLTLTVTVTSGHASTPGEKETATQRSS VPSSTKNAVSMSTSVLSSHSPGSGSTTGQDVTLLAPATPEPASGATWGQDVTSP VTRPALGSTPPPAHDVTSAPDNKPAVGSTAPPAHGVTSAPDTRPAGVTS APDNRPALGSTAPPVHNVTSASGSASGASTLVHNGTSARATTPPASKSTPSPSH SDPTTLASHSTKTDASTHSTVPLTSSNHSSTPQLSTGVSPFLSFHISNDQFNS SLSDPSTDYQELQRDISEMFLQYKQGGFLGSLNIKFRPGSVVQTLAFREGTINV HDVEQFNQYKTEASRNYNLTIQSVSDVPFPFSAQSGAGVGVGIALVLCVLA LAIVYLIALAVCCGCRKNYGGQDLFPARDTYHPMSEPTYHTHGRYVPPSSITRSPYE KEMRVDRERLA"
ORIGIN
Query Match 100.0%; Score 60; DB 12; Length 1414;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTCCACCGCCCCCAGCCCGACGGTGTCACTCGGCCCCGACACCGCGCCCCCG 60
Db 385 GGCTCCACCGCCCCCAGCCCGACGGTGTCACTCGGCCCCGACACCGCGCCCCCG 444
RESULT 4 1455 bp DNA linear PAT 03-FEB-2004
CQ715242
LOCUS Sequence 1176 from Patent WO02068579.
DEFINITION CQ715242
ACCESSION CQ715242
VERSION CQ715242.1 GI:42276099
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..1455
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 60; DB 6; Length 1455;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTCCACCGCCCCCAGCCCGACGGTGTCACTCGGCCCCGACACCGCGCCCCCG 60
Db 412 GGCTCCACCGCCCCCAGCCCGACGGTGTCACTCGGCCCCGACACCGCGCCCCCG 471
RESULT 5 1457 bp DNA linear PAT 14-JAN-2004
AX959914
LOCUS Sequence 19 from Patent WO03100060.
DEFINITION

ACCESSION AX959914
VERSION AX959914.1 GI:40880143
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Burden,N.G., Ellis,J.H. and Hamblin,P.A.
TITLE Muc-1 antigen with reduced number of vnter repeat units
JOURNAL Patent: WO 03100060-A 19 04-DEC-2003;
GLAXO GROUP LIMITED (GB)
FEATURES
SOURCE location/Qualifiers
1..1457
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 60; DB 6; Length 1457;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTCCACCGCGCCCGCAGCCACGGGTGTCACTCGGCGCCCGGACACACGCGCGCCCG 60
|||||
DB 403 GGCTCCACCGCGCCCGCAGCCACGGGTGTCACTCGGCGCCCGGACACACGCGCGCCCG 462
|||||
RESULT 6
AX959684 1774 bp DNA linear PAT 14-JAN-2004
LOCUS
DEFINITION Sequence 10 from Patent WO03099193.
ACCESSION AX959684
VERSION AX959684.1 GI:40880030
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Burden,N. and Hamblin,P.
TITLE Vaccines
JOURNAL Patent: WO 03099193-A 10 04-DEC-2003;
GLAXO GROUP LIMITED (GB)
FEATURES
SOURCE location/Qualifiers
1..1774
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 60; DB 6; Length 1774;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTCCACCGCGCCCGCAGCCACGGGTGTCACTCGGCGCCCGGACACACGCGCGCCCG 60
|||||
DB 462 GGCTCCACCGCGCCCGCAGCCACGGGTGTCACTCGGCGCCCGGACACACGCGCGCCCG 521
|||||
RESULT 7
AX959912 1774 bp DNA linear PAT 14-JAN-2004
LOCUS
DEFINITION Sequence 17 from Patent WO03100060.
ACCESSION AX959912
VERSION AX959912.1 GI:40880142
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Burden,N.G., Ellis,J.H. and Hamblin,P.A.
TITLE Muc-1 antigen with reduced number of vnter repeat units
JOURNAL Patent: WO 03100060-A 17 04-DEC-2003;
GLAXO GROUP LIMITED (GB)
FEATURES
SOURCE location/Qualifiers
1..1774
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 60; DB 6; Length 1774;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTCCACCGCGCCCGCAGCCACGGGTGTCACTCGGCGCCCGGACACACGCGCGCCCG 60
|||||
DB 462 GGCTCCACCGCGCCCGCAGCCACGGGTGTCACTCGGCGCCCGGACACACGCGCGCCCG 521
|||||
RESULT 8
AX922318 1804 bp DNA linear PAT 15-MAY-2004
LOCUS
DEFINITION Sequence 17 from patent US 6716627.
ACCESSION AR4922318
VERSION AR4922318.1 GI:47260892
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1804)
AUTHORS Dobie,K.W.
TITLE Antisense modulation of mucin 1, transmembrane expression
JOURNAL Patent: US 6716627-A 17 06-APR-2004;
FEATURES
SOURCE location/Qualifiers
1..1804
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 60; DB 6; Length 1804;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTCCACCGCGCCCGCAGCCACGGGTGTCACTCGGCGCCCGGACACACGCGCGCCCG 60
|||||
DB 457 GGCTCCACCGCGCCCGCAGCCACGGGTGTCACTCGGCGCCCGGACACACGCGCGCCCG 516
|||||
RESULT 9
AX335367 1804 bp DNA linear PAT 09-JAN-2002
LOCUS
DEFINITION Sequence 5876 from Patent WO0194629.
ACCESSION AX335367
VERSION AX335367.1 GI:18126086
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horrigan,S., Soppet,D.R. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature
JOURNAL Patent: WO 0194629-A 5876 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
SOURCE location/Qualifiers
1..1804
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

TITLE Muc-1 antigen with reduced number of vnter repeat units
JOURNAL Patent: WO 03100060-A 17 04-DEC-2003;
GLAXO GROUP LIMITED (GB)
FEATURES
SOURCE location/Qualifiers
1..1774
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 60; DB 6; Length 1774;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTCCACCGCGCCCGCAGCCACGGGTGTCACTCGGCGCCCGGACACACGCGCGCCCG 60
|||||
DB 462 GGCTCCACCGCGCCCGCAGCCACGGGTGTCACTCGGCGCCCGGACACACGCGCGCCCG 521
|||||
RESULT 8
AR4922318 1804 bp DNA linear PAT 15-MAY-2004
LOCUS
DEFINITION Sequence 17 from patent US 6716627.
ACCESSION AR4922318
VERSION AR4922318.1 GI:47260892
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1804)
AUTHORS Dobie,K.W.
TITLE Antisense modulation of mucin 1, transmembrane expression
JOURNAL Patent: US 6716627-A 17 06-APR-2004;
FEATURES
SOURCE location/Qualifiers
1..1804
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 60; DB 6; Length 1804;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTCCACCGCGCCCGCAGCCACGGGTGTCACTCGGCGCCCGGACACACGCGCGCCCG 60
|||||
DB 457 GGCTCCACCGCGCCCGCAGCCACGGGTGTCACTCGGCGCCCGGACACACGCGCGCCCG 516
|||||
RESULT 9
AX335367 1804 bp DNA linear PAT 09-JAN-2002
LOCUS
DEFINITION Sequence 5876 from Patent WO0194629.
ACCESSION AX335367
VERSION AX335367.1 GI:18126086
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horrigan,S., Soppet,D.R. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature
JOURNAL Patent: WO 0194629-A 5876 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
SOURCE location/Qualifiers
1..1804
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 100.0%; Score 60; DB 6; Length 1804;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GGCTCCACCGCCCCCAGCCACGGTGTACCTCGGCCCCGAGACACGAGCGCCCCG 60
|||||
Db 457 GGCTCCACCGCCCCCAGCCACGGTGTACCTCGGCCCCGAGACACGAGCGCCCCG 516
RESULT 10
HUMMUCAB 1804 bp mRNA linear PRI 07-JAN-1995
LOCUS Human polymorphic epithelial mucin (PEM) mRNA, complete cds.
DEFINITION J05581
ACCESSION J05581
VERSION J05581.1 GI:188869
KEYWORDS polymorphic epithelial mucin.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 1804)
Gendler,S.J., Lancaster,C.A., Taylor-Papadimitriou,J., Duhig,T.,
Peat,N., Burchell,J., Pemberton,L., Lalani,E.N. and Wilson,D.
Molecular cloning and expression of human tumor-associated
polymorphic epithelial mucin
JOURNAL J. Biol. Chem. 265 (25), 15286-15293 (1990)
MEDLINE 90368715
PUBMED 1697589
COMMENT Original source text: Homo sapiens adult adenocarcinoma cDNA to
mRNA.
Draft entry and computer-readable sequence for [J. Biol. Chem.
(1990) in press] kindly submitted
by S.J.Gendler, 26-JUN-1990.
FEATURES
source
1. 1804
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="Xq26.3-q27.1"
/cell_line="BT20"
/tissue_type="adenocarcinoma"
/dev_stage="adult"
1. 1804
/gene="MCNAA"
/product="polymorphic epithelial mucin"
/protein_id="AA59876.1"
/db_xref="GI:188870"
/translation="MTPGTQSPFFLLLLTLVTVTSGHASSTPGGEKETSATORSS
VPSSTEKNAAVSMTSSVLSSHPGSGSSTOGODVTIAPATEPASGSAATWGQDVTSP
VTRPALGSTTPPAHDVTSAPDNKPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTS
APDNRPALGSTAPPAHVNVTASGASGASASTLVNGTSARATTPPASKTSPFSPSH
SDPTTLASHSTKTDASTHSTVPLTSSNHSSTPQLSTGVSEFFLSFHSISNLQFNS
SLEDPSDYYQELQRDISEMFLQIKQGGFLGLSNIKFRPGSVVQQLTLAFREGTINV
HDVETQFNQYKTEAASRYNLTISDVSVDPFPFASGAGVPGWGIALLVLVCVLA
LAIVYLIALAVCQCRKNYQGLDIFPARDTYHPMSEYPTYHTHGRVYPPSSTRSPYE
KVSAGNGSSLSYTNPAVAATSANL"
73. 135
sig_peptide
/gene="MCNAA"
136. 1497
mat_peptide
/gene="MCNAA"
/product="polymorphic epithelial mucin"
1783. 1788
polyA_signal
/gene="MCNAA"
ORIGIN
Query Match 100.0%; Score 60; DB 9; Length 1804;
Best Local Similarity 100.0%; Pred. No. 0.001;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GGCTCCACCGCCCCCAGCCACGGTGTACCTCGGCCCCGAGACACGAGCGCCCCG 60
|||||
Db 457 GGCTCCACCGCCCCCAGCCACGGTGTACCTCGGCCCCGAGACACGAGCGCCCCG 516
RESULT 11
AF423030 1834 bp mRNA linear SYN 10-JUL-2003
LOCUS AF423030
DEFINITION Synthetic construct Homo sapiens mucin variant MUC-1CT80 (MUC1)
ACCESSION AF423030
VERSION AF423030
KEYWORDS AF423030.1 GI:19338619
SOURCE AF423030.1
ORGANISM synthetic construct
synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 1834)
AUTHORS Hinojosa-Kurtzberg,A.M., Johanson,M.E., Madsen,C.S., Hansson,G.C.
and Gendler,S.J.
TITLE Novel MUC1 splice variants contribute to mucin overexpression in
JOURNAL Am. J. Physiol. Gastrointest. Liver Physiol. 284 (5), G853-G862
(2003)
MEDLINE 22570517
PUBMED 12529261
REFERENCE 2 (bases 1 to 1834)
AUTHORS Hinojosa-Kurtzberg,A.M. and Gendler,S.J.
TITLE Direct Submission
JOURNAL Submitted (21-SEP-2001) Biochemistry, Mayo Clinic, 13400 E. Shea
Boulevard, Scottsdale, AZ 85259, USA
FEATURES
source
1. 1834
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/note="Homo sapiens gene in transgenic Mus musculus
C57BL/6; isolated from intestinal mucosa"
1. 1834
/gene="MUC1"
1. 1452
/note="alternatively spliced, contains exon 6a resulting
in variant hydrophilic tail of carboxy-terminal domain;
lacks sites for beta-catenin and Gp2 interactions;
derived from Homo sapiens"
/codon_start=1
/translation="mucin variant MUC-1CT80"
/product="mucin variant MUC-1CT80"
/protein_id="AA186734.1"
/db_xref="GI:19338620"
/translation="MTPGTQSPFFLLLLTLVTVTSGHASSTPGGEKETSATORSS
VPSSTEKNAAVSMTSSVLSSHPGSGSSTOGODVTIAPATEPASGSAATWGQDVTSP
VTRPALGSTTPPAHDVTSAPDNKPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTS
APDNRPALGSTAPPAHVNVTASGASGASASTLVNGTSARATTPPASKTSPFSPSH
SDPTTLASHSTKTDASTHSTVPLTSSNHSSTPQLSTGVSEFFLSFHSISNLQFNS
SLEDPSDYYQELQRDISEMFLQIKQGGFLGLSNIKFRPGSVVQQLTLAFREGTINV
HDVETQFNQYKTEAASRYNLTISDVSVDPFPFASGAGVPGWGIALLVLVCVLA
LAIVYLIALAVCQCRKNYQGLDIFPARDTYHPMSEYPTYHTHGRVYPPSSTRSPYE
KVLRLPTGGQKRWLVLDKDSGGTWKTQRAWKR"
ORIGIN
Query Match 100.0%; Score 60; DB 12; Length 1834;
Best Local Similarity 100.0%; Pred. No. 0.00099;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GGCTCCACCGCCCCCAGCCACGGTGTACCTCGGCCCCGAGACACGAGCGCCCCG 60
|||||
Db 385 GGCTCCACCGCCCCCAGCCACGGTGTACCTCGGCCCCGAGACACGAGCGCCCCG 444
RESULT 12

AX959918
LOCUS AX959918 1835 bp DNA linear PAT 14-JAN-2004
DEFINITION Sequence 23 from Patent WO03100060.
ACCESSION AX959918
VERSION AX959918.1 GI:40880145
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Burden,N.G., Ellis,J.H. and Hamblin,P.A.
TITLE Muc-1 antigen with reduced number of vnter repeat units
JOURNAL Patent: WO 03100060-A 23 04-DEC-2003;
GLAXO GROUP LIMITED (GB)
FEATURES
source location/Qualifiers
1..1835
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 60; DB 6; Length 1835;
Best Local Similarity 100.0%; Pred. No. 0.00099;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTCCACCGCCCCCGACCGGTGTCACTCGGCCCGGACACACCGCGCCCG 60
|||||
Db 462 GGCTCCACCGCCCCCGACCGGTGTCACTCGGCCCGGACACACCGCGCCCG 521
|||||
RESULT 13
AX959916 2135 bp DNA linear PAT 14-JAN-2004
LOCUS AX959916
DEFINITION Sequence 21 from Patent WO03100060.
ACCESSION AX959916
VERSION AX959916.1 GI:40880144
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Burden,N.G., Ellis,J.H. and Hamblin,P.A.
TITLE Muc-1 antigen with reduced number of vnter repeat units
JOURNAL Patent: WO 03100060-A 21 04-DEC-2003;
GLAXO GROUP LIMITED (GB)
FEATURES
source location/Qualifiers
1..2135
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 60; DB 6; Length 2135;
Best Local Similarity 100.0%; Pred. No. 0.00094;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTCCACCGCCCCCGACCGGTGTCACTCGGCCCGGACACACCGCGCCCG 60
|||||
Db 462 GGCTCCACCGCCCCCGACCGGTGTCACTCGGCCCGGACACACCGCGCCCG 521
|||||
RESULT 14
BD272907 2297 bp DNA linear PAT 17-JUL-2003
LOCUS BD272907
DEFINITION A recombinant vector expressing multiple constitutulatory molecules
and uses thereof.
ACCESSION BD272907
VERSION BD272907.1 GI:33082675
KEYWORDS JP 2002531133-A/1.
SOURCE synthetic construct
ORGANISM synthetic construct

artificial sequences.
REFERENCE
AUTHORS 1 (bases 1 to 2297)
TITLE Schlom,J., Hodge,J. and Panicali,D.
JOURNAL A recombinant vector expressing multiple constitutulatory molecules
and uses thereof
PATENT: JP 2002531133-A 1 24-SEP-2002;
COMMENT THE UNITED STATES OF AMERICA, THERION BIOLOGICS CORP
OS Artificial Sequence
PN JP 2002531133-A/1
PD 24-SEP-2002
PF 12-NOV-1999 JP 2000586927
PR 09-DEC-1998 US 60/111582
PI JEFFREY SCHLOM,JAMES HODGE,DENNIS PANICALI
PC C12N15/02,A61K35/12,A61K35/74,A61K35/76,A61K38/00,A61K39/00,
PC A61K39/12,
PC A61K39/125,A61K39/21,A61K39/235,A61K39/245,A61K39/275,A61K39/
PC 29,A61K48/00,
PC A61P1/04,A61P31/04,A61P31/10,A61P31/12,A61P35/00,A61P37/02, PC
A61P37/06,
PC C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N7/00,C12Q1/02,G01N33/
PC 53,C12N15/00,
PC C12N5/00,A61K37/02
CC Description of Artificial Sequence: VECTOR SEQUENCE FH Key
FT source 1..2297
FT location/Qualifiers
1..2297
/organism="Artificial Sequence".
FEATURES
source location/Qualifiers
1..2297
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
ORIGIN
Query Match 100.0%; Score 60; DB 6; Length 2297;
Best Local Similarity 100.0%; Pred. No. 0.00092;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTCCACCGCCCCCGACCGGTGTCACTCGGCCCGGACACACCGCGCCCG 60
|||||
Db 346 GGCTCCACCGCCCCCGACCGGTGTCACTCGGCCCGGACACACCGCGCCCG 405
|||||
RESULT 15
CQ834017 4139 bp DNA linear PAT 29-JUL-2004
LOCUS CQ834017
DEFINITION Sequence 53 from Patent EP1439393.
ACCESSION CQ834017
VERSION CQ834017.1 GI:50833622
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1
Aistle,J.H., Boardman,L.A., Bugart,L.J., Burgess,C.C., Catino,T.J.,
Dwivedi,P., Huntress,M., Johnson,K.A., Lewis,M.E., Maimonis,P.J.,
Myerow,S.H., Brown-Shimer,S.L., Thiagalingam,A., Thibodeau,S.N. and
Molino,G.A.
TITLE Detection methods using TIMP 1 for colon cancer diagnosis
JOURNAL Patent: EP 1439393-A 53 21-JUL-2004;
Bayer Healthcare LLC (US); MAYO FOUNDATION FOR MEDICAL EDUCATION
AND RESEARCH (US)
FEATURES
source location/Qualifiers
1..4139
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 60; DB 6; Length 4139;
Best Local Similarity 100.0%; Pred. No. 0.00076;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Sun Jan 16 16:35:03 2005

us-10-057-136-2.rge

Page 6

OY 1 GGCTCCACCGCCCCCCCAGGCCACGGTGTCACTCGGCCCCGACAACGCGGCCTCC 60
| | | | |
Dlb 458 GGCATCACCGCCCCCAGGCCACGGTGTCACCTCGGCCCCGACAACGAGCGGCCTCC 517
| | | | |

Search completed: January 15, 2005, 22:36:42
Job time : 724.8 secs

Db 1 GGCTCCACCGCCCCCGCCAGCCACGGTGTCACCTCGGCCCGGACACACAGCGCGGCCCG 60

RESULT 2

US-10-447-839A-75/c
; Sequence 75, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald W.
; APPLICANT: Kharbanda, Surrender
; APPLICANT: Weitman, Steven D.
; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT APPLICATION NUMBER: US/10/447, 839A
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 75
; LENGTH: 1424
; TYPE: RNA
; ORGANISM: ARTIFICIAL
; FEATURE:
; OTHER INFORMATION: Synthesized Sequence
US-10-447-839A-75

Query Match 100.0%; Score 60; DB 16; Length 1424;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCGCCAGCCACGGTGTCACCTCGGCCCGGACACACAGCGCGGCCCG 60
Db 1041 GGCTCCACCGCCCCCGCCAGCCACGGTGTCACCTCGGCCCGGACACACAGCGCGGCCCG 982

RESULT 3

US-10-447-839A-20
; Sequence 20, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald W.
; APPLICANT: Kharbanda, Surrender
; APPLICANT: Weitman, Steven D.
; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT APPLICATION NUMBER: US/10/447, 839A
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 1428
; TYPE: RNA
; ORGANISM: RNA
US-10-447-839A-20

Query Match 100.0%; Score 60; DB 16; Length 1428;
Best Local Similarity 93.3%; Pred. No. 1.3e-09;
Matches 56; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCGCCAGCCACGGTGTCACCTCGGCCCGGACACACAGCGCGGCCCG 60
Db 385 GGCTCCACCGCCCCCGCCAGCCACGGTGTCACCTCGGCCCGGACACACAGCGCGGCCCG 444

RESULT 4

US-10-057-136-19
; Sequence 19, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUF, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 1527
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-19

Query Match 100.0%; Score 60; DB 14; Length 1527;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCGCCAGCCACGGTGTCACCTCGGCCCGGACACACAGCGCGGCCCG 60
Db 226 GGCTCCACCGCCCCCGCCAGCCACGGTGTCACCTCGGCCCGGACACACAGCGCGGCCCG 285

RESULT 5

US-10-447-839A-19
; Sequence 19, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald W.
; APPLICANT: Kharbanda, Surrender
; APPLICANT: Weitman, Steven D.
; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT APPLICATION NUMBER: US/10/447, 839A
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 1799
; TYPE: RNA
; ORGANISM: RNA
US-10-447-839A-19

Query Match 100.0%; Score 60; DB 16; Length 1799;
Best Local Similarity 93.3%; Pred. No. 1.2e-09;
Matches 56; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCGCCAGCCACGGTGTCACCTCGGCCCGGACACACAGCGCGGCCCG 60
Db 453 GGCTCCACCGCCCCCGCCAGCCACGGTGTCACCTCGGCCCGGACACACAGCGCGGCCCG 512

RESULT 6

US-09-964-824A-573
; Sequence 573, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horriagan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964, 824A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236, 033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236, 032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236, 028
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 573
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-824A-573

Query Match 100.0%; Score 60; DB 9; Length 1804;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCGCCAGCCACGGTGTCACTCGGCGCCCGGACACACAGCGCCGCCCG 60
|||||
Db 457 GGCTCCACCGCCCGCCAGCCACGGTGTCACTCGGCGCCCGGACACACAGCGCCGCCCG 516

RESULT 7

US-10-029-517-17
; Sequence 17, Application US/10029517
; Publication No. US20030148969A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029, 517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 17
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (73)...(1500)
US-10-029-517-17

Query Match 100.0%; Score 60; DB 15; Length 1804;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCGCCAGCCACGGTGTCACTCGGCGCCCGGACACACAGCGCCGCCCG 60
|||||
Db 457 GGCTCCACCGCCCGCCAGCCACGGTGTCACTCGGCGCCCGGACACACAGCGCCGCCCG 516

RESULT 8

US-10-717-597-30
; Sequence 30, Application US/10717597
; Publication No. US20040110221A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Butczynski, Michael E.

; APPLICANT: Twine, Natalie C.
; APPLICANT: Dörner, Andrew J.
; APPLICANT: Trepicchio, William L.
; APPLICANT: Slonim, Donna K.
; APPLICANT: Stover, Jennifer A.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING RCC AND OTHER SOLID TUMORS
; FILE REFERENCE: AM101080L
; CURRENT APPLICATION NUMBER: US/10/717, 597
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US 60/459, 782
; PRIOR FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: US 60/427, 982
; PRIOR FILING DATE: 2002-11-21
; NUMBER OF SEQ ID NOS: 4904
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-717-597-30

Query Match 100.0%; Score 60; DB 17; Length 1804;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCGCCAGCCACGGTGTCACTCGGCGCCCGGACACACAGCGCCGCCCG 60
|||||
Db 457 GGCTCCACCGCCCGCCAGCCACGGTGTCACTCGGCGCCCGGACACACAGCGCCGCCCG 516

RESULT 9

US-10-775-920-84
; Sequence 84, Application US/10775920
; Publication No. US20040175744A1
; GENERAL INFORMATION:
; APPLICANT: Mergen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; FILE REFERENCE: Mergen - 0010B
; CURRENT APPLICATION NUMBER: US/10/775, 920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447, 900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 84
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-84

Query Match 100.0%; Score 60; DB 17; Length 1804;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCGCCAGCCACGGTGTCACTCGGCGCCCGGACACACAGCGCCGCCCG 60
|||||
Db 457 GGCTCCACCGCCCGCCAGCCACGGTGTCACTCGGCGCCCGGACACACAGCGCCGCCCG 516

RESULT 10

US-10-101-510-339
; Sequence 339, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101, 510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276, 947

```
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 339
; LENGTH: 1823
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-339
```

```
Query Match          100.0%; Score 60; DB 15; Length 1823;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GGCTCCACCGCCCCCGCCAGCCAGCGTGTCACTCGGCCCCCGGACACCAAGCGCGCCCCG 60
          |||
Db      457 GGCTCCACCGCCCCCGCCAGCCAGCGTGTCACTCGGCCCCCGGACACCAAGCGCGCCCCG 516
```

RESULT 11

```
US-10-406-317-41
; Sequence 41, Application US/10406317
; Publication No. US20040019195A1
; GENERAL INFORMATION:
; APPLICANT: Schlom, Jeffrey;
; APPLICANT: Hodge, James;
; APPLICANT: Panicali, Dennis
; TITLE OF INVENTION: A recombinant vector expressing multiple constitutulatory
; FILE REFERENCE: 38163-0189
; CURRENT FILING DATE: 2003-04-04
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: US/09/856,988
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: PCT/US99/26866
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 60/111,582
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: VECTOR
US-10-406-317-41
```

```
Query Match          100.0%; Score 60; DB 16; Length 2297;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GGCTCCACCGCCCCCGCCAGCGTGTCACTCGGCCCCCGGACACCAAGCGCGCCCCG 60
          |||
Db      346 GGCTCCACCGCCCCCGCCAGCCAGCGTGTCACTCGGCCCCCGGACACCAAGCGCGCCCCG 405
```

RESULT 12

```
US-09-964-824A-105
; Sequence 105, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horrigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964,824A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
```

```
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 105
; LENGTH: 4139
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-824A-105
```

```
Query Match          100.0%; Score 60; DB 9; Length 4139;
Best Local Similarity 100.0%; Pred. No. 9.7e-10;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GGCTCCACCGCCCCCGCCAGCCAGCGTGTCACTCGGCCCCCGGACACCAAGCGCGCCCCG 60
          |||
Db      458 GGCTCCACCGCCCCCGCCAGCCAGCGTGTCACTCGGCCCCCGGACACCAAGCGCGCCCCG 517
```

RESULT 13

```
US-09-964-824A-578
; Sequence 578, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horrigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964,824A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 578
; LENGTH: 4139
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-824A-578
```

```
Query Match          100.0%; Score 60; DB 9; Length 4139;
Best Local Similarity 100.0%; Pred. No. 9.7e-10;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GGCTCCACCGCCCCCGCCAGCCAGCGTGTCACTCGGCCCCCGGACACCAAGCGCGCCCCG 60
          |||
Db      458 GGCTCCACCGCCCCCGCCAGCCAGCGTGTCACTCGGCCCCCGGACACCAAGCGCGCCCCG 517
```

RESULT 14

```
US-09-864-864-334
; Sequence 334, Application US/09864864
; Patent No. US20020102679A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Secrist, Heather
; APPLICANT: Lodes, Michael J.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steve P.
; APPLICANT: Mannion, Jane
; APPLICANT: Benson, Darin R.
; APPLICANT: Carter, Darick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.523
; CURRENT APPLICATION NUMBER: US/09/864,864
; CURRENT FILING DATE: 2001-05-23
```

```
; NUMBER OF SEQ ID NOS: 341
; SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 334
; LENGTH: 4139
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-864-334
```

```
Query Match          100.0%; Score 60; DB 9; Length 4139;
Best Local Similarity 100.0%; Pred. No. 9.7e-10;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GGCCTCACCGCCCCCGACCGGTGTCACTCGGCCCCGACACACGCGCCCGG 60
         |||||||
Db      458 GGCCTCACCGCCCCCGACCGGTGTCACTCGGCCCCGACACACGCGCCCGG 517
```

RESULT 15

US-09-880-107-2121

```
; Sequence 2121, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2121
; LENGTH: 4139
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 J05582
US-09-880-107-2121
```

```
Query Match          100.0%; Score 60; DB 9; Length 4139;
Best Local Similarity 100.0%; Pred. No. 9.7e-10;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GGCCTCACCGCCCCCGACCGGTGTCACTCGGCCCCGACACACGCGCCCGG 60
         |||||||
Db      458 GGCCTCACCGCCCCCGACCGGTGTCACTCGGCCCCGACACACGCGCCCGG 517
```

Search completed: January 16, 2005, 09:30:18
Job time : 184.4 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: January 15, 2005, 19:50:13 ; Search time 35.2 Seconds
(without alignments)
1211.572 Million cell updates/sec

Title: US-10-057-136-2

Perfect score: 60
Sequence: 1 ggcctcaccgccccccagc.....cggacaccagcgccgccccg 60

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	100.0	1804	4 US-10-029-517-17	Sequence 17, Appl
2	60	100.0	8186	4 US-10-029-517-19	Sequence 19, Appl
3	58.4	97.3	572	4 US-10-029-517-18	Sequence 18, Appl
4	58.4	97.3	1721	4 US-10-029-517-3	Sequence 3, Appl
5	57	95.0	518	4 US-10-029-517-101	Sequence 101, App
6	55.2	92.0	3343	4 US-10-029-517-102	Sequence 102, App
7	54	90.0	60	4 US-09-475-947A-246	Sequence 246, App
8	52.2	87.0	981	4 US-10-029-517-16	Sequence 16, Appl
9	51	85.0	6192	2 US-08-479-537A-1	Sequence 1, Appl
10	51	85.0	6192	3 US-09-083-116-1	Sequence 1, Appl
11	51	85.0	6192	3 US-09-134-916A-1	Sequence 1, Appl
12	51	85.0	6449	2 US-08-479-537A-4	Sequence 4, Appl
13	51	85.0	6449	3 US-09-083-116-4	Sequence 4, Appl
14	51	85.0	6449	3 US-09-134-916A-4	Sequence 4, Appl
15	29.6	49.3	556	4 US-10-029-517-105	Sequence 105, App
16	27.6	46.0	5577	1 US-08-326-117B-1	Sequence 1, Appl
17	27.6	46.0	5577	3 US-08-982-129-1	Sequence 1, Appl
18	27.6	46.0	5582	4 US-09-178-176B-1	Sequence 1, Appl
19	27.6	46.0	5582	4 US-09-457-864-1	Sequence 1, Appl
20	27.6	46.0	5582	4 US-09-457-865A-1	Sequence 1, Appl
21	27.2	45.3	1441	1 US-08-136-277-18	Sequence 18, Appl
22	27.2	45.3	1441	2 US-08-479-403-18	Sequence 18, Appl
23	27.2	45.3	1441	3 US-08-835-734-18	Sequence 18, Appl
24	27.2	45.3	2750	1 US-08-136-277-1	Sequence 1, Appl
25	27.2	45.3	2750	2 US-08-479-403-1	Sequence 1, Appl
26	27.2	45.3	2750	3 US-08-835-734-1	Sequence 1, Appl
27	27.2	45.3	77536	4 US-09-410-551B-1	Sequence 1, Appl

28	27.2	45.3	77536	4 US-09-940-316B-1	Sequence 1, Appl
C 29	27	45.0	423	4 US-09-854-133-713	Sequence 713, App
30	26.8	44.7	519	4 US-09-646-028-42	Sequence 42, Appl
31	26.8	44.7	534	4 US-09-646-028-46	Sequence 46, Appl
C 32	26.6	44.3	1737	2 US-08-750-703-2	Sequence 2, Appl
C 33	26.6	44.3	8779	2 US-08-750-703-4	Sequence 4, Appl
C 34	26.4	44.0	774	3 US-08-415-655-11	Sequence 11, Appl
C 35	26.4	44.0	3245	4 US-09-774-528-352	Sequence 352, App
C 36	26.4	44.0	4480	3 US-09-191-171-7	Sequence 7, Appl
C 37	26.4	44.0	4480	3 US-09-385-707-7	Sequence 7, Appl
C 38	26.4	44.0	6238	4 US-09-639-696C-6	Sequence 6, Appl
C 39	26.4	44.0	8438	1 US-07-785-420-1	Sequence 1, Appl
C 40	26.4	44.0	15378	3 US-08-785-420-1	Sequence 1, Appl
C 41	26	43.3	2581	3 US-09-363-708-1	Sequence 1, Appl
C 42	26	43.3	2581	4 US-09-083-587-1	Sequence 1, Appl
C 43	26	43.3	2743	1 US-08-818-823-3	Sequence 3, Appl
44	26	43.3	2743	1 US-08-818-823-3	Sequence 3, Appl
45	26	43.3	2749	2 US-08-124-981A-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-10-029-517-17
; Sequence 17, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 17
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (73)...(1500)
US-10-029-517-17

Query Match 100.0%; Score 60; DB 4; Length 1804;
Best Local Similarity 100.0%; Pred. No. 6e-08;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTCACCAGCCCGCCAGCCAGCGGTGTACCTCGGCCCCGAGACAGCGCGGCCG 60
Db 457 GGCCTCACCAGCCCGCCCGCCAGCCAGCGGTGTACCTCGGCCCCGAGACAGCGCGGCCG 516

RESULT 2
US-10-029-517-19
; Sequence 19, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 19
; LENGTH: 8186
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 6899
; OTHER INFORMATION: unknown


```
/ NAME/KEY: unsure
/ LOCATION: 7155
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 7184
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 7957
/ OTHER INFORMATION: unknown
/ NAME/KEY: intron
/ LOCATION: (2997)...(3498)
/ OTHER INFORMATION: intron 1
/ NAME/KEY: intron:exon junction
/ LOCATION: (3498)...(3499)
/ OTHER INFORMATION: intron 1:exon 2
/ NAME/KEY: exon
/ LOCATION: (3508)...(3599)
/ OTHER INFORMATION: exon 2d
/ NAME/KEY: exon:intron junction
/ LOCATION: (3982)...(3983)
/ OTHER INFORMATION: exon 2a:intron 2a
/ NAME/KEY: intron:exon junction
/ LOCATION: (4205)...(4206)
/ OTHER INFORMATION: intron 2c:exon 3c
/ NAME/KEY: intron:exon junction
/ LOCATION: (4259)...(4260)
/ OTHER INFORMATION: intron 2d:exon 3d
/ NAME/KEY: exon
/ LOCATION: (4260)...(4328)
/ OTHER INFORMATION: exon 3d
/ NAME/KEY: intron:exon junction
/ LOCATION: (4632)...(4633)
/ OTHER INFORMATION: intron 3:exon 4
/ NAME/KEY: exon
/ LOCATION: (4914)...(5035)
/ OTHER INFORMATION: exon 5
/ NAME/KEY: intron
/ LOCATION: (5266)...(6293)
/ OTHER INFORMATION: intron 6
/ US-10-029-517-19

Query Match
Best Local Similarity 100.0%; Score 60; DB 4; Length 8186;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCGCCAGCGGTGTCACTCGGCCCCGAGACACGAGCGGCCCG 60
DB 3825 GGCTCCACCGCCCCCGCCAGCGGTGTCACTCGGCCCCGAGACACGAGCGGCCCG 3884

RESULT 3
US-10-029-517-18
/ Sequence 18, Application US/10029517
/ Patent No. 6716627
/ GENERAL INFORMATION:
/ APPLICANT: Kenneth W. Dobie
/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
/ FILE REFERENCE: RTS-0352
/ CURRENT APPLICATION NUMBER: US/10/029,517
/ CURRENT FILING DATE: 2001-12-20
/ NUMBER OF SEQ ID NOS: 107
/ SEQ ID NO 18
/ LENGTH: 572
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (67)...(572)
/ US-10-029-517-18

Query Match
Best Local Similarity 97.3%; Score 58.4; DB 4; Length 572;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCGCCAGCGGTGTCACTCGGCCCCGAGACACGAGCGGCCCG 60
DB 478 GGCTCCACCGCCCCCGCCAGCGGTGTCACTCGGCCCCGAGACACGAGCGGCCCG 537

RESULT 4
US-10-029-517-3
/ Sequence 3, Application US/10029517
/ Patent No. 6716627
/ GENERAL INFORMATION:
/ APPLICANT: Kenneth W. Dobie
/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
/ FILE REFERENCE: RTS-0352
/ CURRENT APPLICATION NUMBER: US/10/029,517
/ CURRENT FILING DATE: 2001-12-20
/ NUMBER OF SEQ ID NOS: 107
/ SEQ ID NO 3
/ LENGTH: 1721
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (58)...(1605)
/ US-10-029-517-3

Query Match
Best Local Similarity 97.3%; Score 58.4; DB 4; Length 1721;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCGCCAGCGGTGTCACTCGGCCCCGAGACACGAGCGGCCCG 60
DB 442 GGCTCCACCGCCCCCGCCAGCGGTGTCACTCGGCCCCGAGACACGAGCGGCCCG 501

RESULT 5
US-10-029-517-101
/ Sequence 101, Application US/10029517
/ Patent No. 6716627
/ GENERAL INFORMATION:
/ APPLICANT: Kenneth W. Dobie
/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
/ FILE REFERENCE: RTS-0352
/ CURRENT APPLICATION NUMBER: US/10/029,517
/ CURRENT FILING DATE: 2001-12-20
/ NUMBER OF SEQ ID NOS: 107
/ SEQ ID NO 101
/ LENGTH: 518
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ US-10-029-517-101

Query Match
Best Local Similarity 95.0%; Score 57; DB 4; Length 518;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCGCCAGCGGTGTCACTCGGCCCCGAGACACGAGCGGCC 57
DB 462 GGCTCCACCGCCCCCGCCAGCGGTGTCACTCGGCCCCGAGACACGAGCGGCC 518

RESULT 6
US-10-029-517-102
/ Sequence 102, Application US/10029517
/ Patent No. 6716627
/ GENERAL INFORMATION:
/ APPLICANT: Kenneth W. Dobie
/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
```

```
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 102
; LENGTH: 3343
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-10-029-517-102
```

```
Query Match          92.0%; Score 55.2; DB 4; Length 3343;
Best Local Similarity 95.0%; Pred. No. 1.1e-06;
Matches 57; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY      1 GGCCTCACCGCCCGCCAGCCAGCGTGTCACTCGGCGCCCGGACACACGAGCCGCGCCG 60
          |||||||
Db       1668 GGCCTCACCGCCCGCCAGCCAGCGTGTCACTCGGCGCCCGGAGACACGAGCCGCGCCCG 1727
```

```
RESULT 7
US-09-475-947A-246
; Sequence 246, Application US/09475947A
; Patent No. 6472154
; GENERAL INFORMATION:
; APPLICANT: Garner, Harold R.
; APPLICANT: Wren, Jonathan D.
; APPLICANT: Mina, John D.
; TITLE OF INVENTION: Polymorphic Repeats in Human Genes
; FILE REFERENCE: UTS0667
; CURRENT APPLICATION NUMBER: US/09/475,947A
; CURRENT FILING DATE: 1999-12-31
; NUMBER OF SEQ ID NOS: 346
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 246
; LENGTH: 60
; TYPE: DNA
; ORGANISM: human
US-09-475-947A-246
```

Query Match

```
Best Local Similarity 100.0%; Score 54; DB 4; Length 60;
Pred. No. 2.8e-06;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GGCCTCACCGCCCGCCAGCCAGCGTGTCACTCGGCGCCCGGACACACGAGCCG 54
          |||||||
Db       7 GGCCTCACCGCCCGCCAGCCAGCGTGTCACTCGGCGCCCGGACACACGAGCCG 60
```

RESULT 8

```
US-10-029-517-16
; Sequence 16, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 16
; LENGTH: 981
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon:exon junction
; LOCATION: (464)...(465)
; OTHER INFORMATION: exon 3b:exon 4
US-10-029-517-16
```

```
Query Match          87.0%; Score 52.2; DB 4; Length 981;
Best Local Similarity 94.7%; Pred. No. 7.3e-06;
```

```
Matches 54; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY      1 GGCCTCACCGCCCGCCAGCCAGCGTGTCACTCGGCGCCCGGACACACGAGCCGCGCC 57
          |||||||
Db       21 GGCCTCACCGCCCGCCAGCCAGCGTGTCACTCGGCGCCCGGACACACGAGCCGCGCC 77
```

RESULT 9

```
US-08-479-537A-1
; Sequence 1, Application US/08479537A
; Patent No. 5861381
; GENERAL INFORMATION:
; APPLICANT: CHAMON, Pierre
; APPLICANT: KIENY, Marie-Paule
; APPLICANT: LATHE, Richard
; APPLICANT: HAREUVENI, Mara
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
; TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,537A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 90/13101
; FILING DATE: 23-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR91/00835
; FILING DATE: 23-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/039,320
; FILING DATE: 04-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,576
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teekin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 017753-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6192 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 58..120
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 439..5239
; OTHER INFORMATION: /note= "The nucleotides spanning
; OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
; OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
; OTHER INFORMATION: The number of such repeats varies from 1 to 80."
; FEATURE:
; NAME/KEY: mat_peptide
```

```
/ LOCATION: 121..6166
/ FEATURE:
/ NAME/KEY: repeat_region
/ LOCATION: 457
/ OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
/ OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
/ OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
/ FEATURE:
/ NAME/KEY: repeat_region
/ LOCATION: 487
/ OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
/ OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
/ OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
/ FEATURE:
/ NAME/KEY: repeat_region
/ LOCATION: 496
/ OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
/ OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
/ OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-08-479-537A-1
```

```
Query Match 85.0%; Score 51; DB 2; Length 6192;
Best Local Similarity 85.0%; Pred. No. 1.4e-05;
Matches 51; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

```
Qy 1 GGCTCCACCGCCCCCGACCGCGTGTCACTCGGCCCCGACACACCGCGCCCCG 60
DB 442 GGCTCCACCGCCCCCGACCGCGTGTCACTCGGCCCCGACACACCGCGCCCCG 501
```

RESULT 10

```
US-09-083-116-1
; Sequence 1, Application US/09083116
; Patent No. 6203795
; GENERAL INFORMATION:
; APPLICANT: CHAMBER, Pierre
; APPLICANT: KIENY, Marie-Paule
; APPLICANT: LATHE, Richard
; APPLICANT: HAREUVENI, Mara
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
; TITLE OF INVENTION: TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/083,116
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,537
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR91/00835
; FILING DATE: 23-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/039,320
; FILING DATE: 04-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,576
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
```

```
REFERENCE/DOCKET NUMBER: 017753-025
```

```
TELECOMMUNICATION INFORMATION:
```

```
TELEPHONE: (703) 836-6620
```

```
TELEFAX: (703) 836-2021
```

```
INFORMATION FOR SEQ ID NO: 1:
```

```
SEQUENCE CHARACTERISTICS:
```

```
LENGTH: 6192 base pairs
```

```
TYPE: nucleic acid
```

```
STRANDEDNESS: single
```

```
TOPOLOGY: linear
```

```
MOLECULE TYPE: DNA (genomic)
```

```
FEATURE:
```

```
NAME/KEY: sig_peptide
```

```
LOCATION: 58..120
```

```
FEATURE:
```

```
NAME/KEY: repeat_region
```

```
LOCATION: 439..5239
```

```
OTHER INFORMATION: /note= "The nucleotides spanning
```

```
OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
```

```
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
```

```
OTHER INFORMATION: The number of such repeats varies from 1 to 80."
```

```
FEATURE:
```

```
NAME/KEY: mat_peptide
```

```
LOCATION: 121..6166
```

```
FEATURE:
```

```
NAME/KEY: repeat_region
```

```
LOCATION: 457
```

```
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
```

```
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
```

```
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
```

```
FEATURE:
```

```
NAME/KEY: repeat_region
```

```
LOCATION: 487
```

```
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
```

```
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
```

```
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
```

```
FEATURE:
```

```
NAME/KEY: repeat_region
```

```
LOCATION: 496
```

```
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
```

```
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
```

```
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
```

```
US-09-083-116-1
```

```
Query Match 85.0%; Score 51; DB 3; Length 6192;
```

```
Best Local Similarity 85.0%; Pred. No. 1.4e-05;
```

```
Matches 51; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

```
Qy 1 GGCTCCACCGCCCCCGACCGCGTGTCACTCGGCCCCGACACACCGCGCCCCG 60
DB 442 GGCTCCACCGCCCCCGACCGCGTGTCACTCGGCCCCGACACACCGCGCCCCG 501
```

RESULT 11

```
US-09-134-916A-1
; Sequence 1, Application US/09134916A
; Patent No. 6328956
; GENERAL INFORMATION:
; APPLICANT: CHAMBER, Pierre
; APPLICANT: KIENY, Marie-Paule
; APPLICANT: LATHE, Richard
; APPLICANT: HAREUVENI, Mara
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
; TITLE OF INVENTION: TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
```

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/134,916A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/479,537
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: FR 90/13101
FILING DATE: 23-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teekln, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6192 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note= "The nucleotides spanning
OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
OTHER INFORMATION: The number of such repeats varies from 1 to 80."
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 121..6166
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-09-134-916A-1
Query Match 85.0%; Score 51; DB 3; Length 6192;
Best Local Similarity 85.0%; Pred. No. 1.4e-05;
Matches 51; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCAGCCCGGTGTACCTCGCCCCCGACACCGCGCCCCG 60
|||||
Db 442 GGCTCCACCGCCCCCANNNGCCCGGTGTACCTCGCCCCCGACANNAGCCGNNCCG 501
|||||
RESULT 12
US-08-479-537A-4
Sequence 4, Application US/08479537A
Patent No. 5861381
GENERAL INFORMATION:
APPLICANT: CHAMON, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREUVENI, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/479,537A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 90/13101
FILING DATE: 23-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teekln, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6449 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note= "The nucleotides spanning
OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
OTHER INFORMATION: The number of such repeats varies from 1 to 80."
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 121..5661
FEATURE:

NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-08-479-537A-4

Query Match 85.0%; Score 51; DB 2; Length 6449;
Best Local Similarity 85.0%; Pred. No. 1.4e-05;
Matches 51; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCGCCAGCCGCGTGTACCTCGCCCGGACACACAGCGCCCGG 60
Db 442 GGCTCCACCGCCCCCGCCAGCCGCGTGTACCTCGCCCGGACACACAGCGCCCGG 501

RESULT 13

US-09-083-116-4
Sequence 4, Application US/09083116
Patent No. 6203795

GENERAL INFORMATION:
APPLICANT: CHAMON, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREUVENI, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/083,116
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,537
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6449 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note= "The nucleotides spanning
OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
OTHER INFORMATION: The number of such repeats varies from 1 to 80."
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 121..5661
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, AC
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."

US-09-083-116-4

Query Match 85.0%; Score 51; DB 3; Length 6449;
Best Local Similarity 85.0%; Pred. No. 1.4e-05;
Matches 51; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCGCCAGCCGCGTGTACCTCGCCCGGACACACAGCGCCCGG 60
Db 442 GGCTCCACCGCCCCCGCCAGCCGCGTGTACCTCGCCCGGACACACAGCGCCCGG 501

RESULT 14

US-09-134-916A-4
Sequence 4, Application US/09134916A
Patent No. 6328956

GENERAL INFORMATION:
APPLICANT: CHAMON, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREUVENI, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/134,916A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/479,537
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: FR 90/13101
FILING DATE: 23-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6449 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note= "The nucleotides spanning
OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
OTHER INFORMATION: The number of such repeats varies from 1 to 80."
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 121..5661
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-09-134-916A-4
```

```
Query Match      85.0%; Score 51; DB 3; Length 6449;
Best Local Similarity 85.0%; Pred. No. 1.4e-05;
Matches 51; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

```
1 GGCTCCACCGCCCCCGACCGGTCACCTCGGCGCCGACACACGCGCGCCCG 60
|||||
```

```
Db      442 GGCTCCACCGCCCCCGACCGGTCACCTCGGCGCCGACACGCGCGCCCG 501
RESULT 15
US-10-029-517-105
; Sequence 105, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 105
; LENGTH: 556
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5
; OTHER INFORMATION: n = A,T,C or G
US-10-029-517-105
```

```
Query Match      49.3%; Score 29.6; DB 4; Length 556;
Best Local Similarity 68.3%; Pred. No. 7.4;
Matches 41; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
```

```
QY      1 GGCTCCACCGCCCCCGACCGGTCACCTCGGCGCCGACACGCGCGCCCG 60
Db      481 GGTTCAAGCTGCACCTGGGACAGGATGTCACTTCCTCCAGTCACGAGGCCAGCCCTG 540
|||||
```

Search completed: January 16, 2005, 03:01:14
Job time : 36.2 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 16:00:11 ; Search time 170.5 Seconds
(without alignments)
1847.304 Million cell updates/sec

Title: US-10-057-136-2

Perfect score: 60
Sequence: 1 ggcaccacgcgcccccagc.....cggacacaggcgcccg 60

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 beqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum	DB seq	length:	0
Maximum	DB seq	length:	0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N_Geneseq_23Sep04:*

```

1: geneseqn1980s: *
2: geneseqn1990s: *
3: geneseqn2000s: *
4: geneseqn2001as: *
5: geneseqn2001bs: *
6: geneseqn2002as: *
7: geneseqn2002bs: *
8: geneseqn2003as: *
9: geneseqn2003bs: *
10: geneseqn2003cs: *
11: geneseqn2003ds: *
12: geneseqn2004s: *

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	60	100.0	60	2	AAV48316	Aav48316 Nucleotide
2	60	100.0	309	1	AAN90579	Aan90579 pDF9.3 cD
3	60	100.0	1194	12	ADIS7712	Adis7712 Human bre
4	60	100.0	1378	12	ADIS7653	Adis7653 Human bre
5	60	100.0	1424	12	ADO23180	Ado23180 Antisense
6	60	100.0	1428	6	ABL60159	Ab160159 Human MUC
7	60	100.0	1428	12	ADO23125	Ado23125 Human MUC
8	60	100.0	1457	12	ADF32627	Adf32627 Plasmid J
9	60	100.0	1527	2	AAV48329	Aav48329 MiniMUC1
10	60	100.0	1614	12	ADK70370	Adk70370 Respiratc
11	60	100.0	1630	12	ADIS7708	Adis7708 Human bre
12	60	100.0	1634	12	ADIS7689	Adis7689 Human bre
13	60	100.0	1712	12	ADIS7666	Adis7666 Human bre
14	60	100.0	1738	12	ADIS7669	Adis7669 Human bre
15	60	100.0	1755	12	ADIS7673	Adis7673 Human bre
16	60	100.0	1774	12	ADE43991	Ade43991 Plasmid J
17	60	100.0	1774	12	ADf32625	Adf32625 Plasmid J
18	60	100.0	1799	12	ADO23124	Ado23124 Human MUC
19	60	100.0	1803	12	ADIS7699	Adis7699 Human bre
20	60	100.0	1804	6	ABL67539	Ab167539 Thyroid c
21	60	100.0	1804	9	AAD56950	Aad56950 Human muc

22	60	100.0	1804	10	ADD14719	Ad14719	Human	src
23	60	100.0	1804	12	ADP13294	Adp13294	Renal	cel
24	60	100.0	1805	12	ADO28642	Ado28642	Human	MUC
25	60	100.0	1808	12	AD157706	Ad157706	Human	bre
26	60	100.0	1818	12	ADF32633	Adf32633	Plaemid	J
27	60	100.0	1823	6	ABZ35228	Abz35228	Human	gen
28	60	100.0	1823	12	AD157707	Ad157707	Human	bre
29	60	100.0	1835	12	ADF32631	Adf32631	Plaemid	J
30	60	100.0	1874	12	AD157688	Ad157688	Human	bre
31	60	100.0	1882	12	AD157677	Ad157677	Human	bre
32	60	100.0	1918	12	AD157672	Ad157672	Human	bre
33	60	100.0	1930	12	AD157678	Ad157678	Human	bre
34	60	100.0	1945	12	AD157676	Ad157676	Human	bre
35	60	100.0	1949	12	AD157698	Ad157698	Human	bre
36	60	100.0	1953	12	AD157668	Ad157668	Human	bre
37	60	100.0	2045	12	AD157701	Ad157701	Human	bre
38	60	100.0	2049	12	AD157682	Ad157682	Human	bre
39	60	100.0	2090	12	AD157705	Ad157705	Human	bre
40	60	100.0	2094	12	AD157681	Ad157681	Human	bre
41	60	100.0	2135	12	ADF32629	Adf32629	Plaemid	J
42	60	100.0	2194	12	AD157683	Ad157683	Human	bre
43	60	100.0	2255	12	AD157667	Ad157667	Human	bre
44	60	100.0	2333	12	AD157685	Ad157685	Human	bre
45	60	100.0	4139	6	AB167071	Ab167071	Thyroid	c

ALIGNMENTS

RESULT 1

ID AAV48316 standard; cDNA; 60 BP.

AC AAV48316;

DT 20-NOV-1998 (first entry)

DE Nucleotide sequence encoding MUC1 tandem repeat unit.

KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;

2000

[illegible]

FT	CDS	1.	.60
----	-----	----	-----

```

FT      /product= "MUC1 tandem repeat"
....

```

PN W09837095-A2.

PD 27-AUG-1998.

PF 24-FEB-1998; 98WO-US003693.

PR 24-FEB-1997; 97US-0038253P.

PA (THER-) THERION BIOLOGICS CORP.

PA (DAND) DANA FARBER CANCER INST INC.

PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;

DR WPI; 1998-467492/40.

XX
XX
XX

PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.

Example 1; Page 20; 42pp; English.

CC The MUC1 tandem repeat unit was used to create an immunogenic mini-MUC1

CC in a pharmaceutical composition also containing an immunomodulator to
CC generate MUC1 specific cytotoxic T-lymphocytes. The recombinant pox virus
CC therefore encodes an immunogenic MUC1 fragment that does not undergo
CC significant genetic deletion, thereby providing an unexpectedly stable
CC and immunogenic pox virus. They can be used to prevent or treat tumours
XX expressing MUC1 tumour-associated antigens
SQ Sequence 60 BP; 7 A; 33 C; 16 G; 4 T; 0 U; 0 Other;

Query Match
Best Local Similarity 100.0%; Score 60; DB 2; Length 60;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCTCCACCGCGCCCGCCAGCCCGGTGTCACTCGGCCCGGACACACGCGCCCGCG 60
Db 1 GGCTCCACCGCGCCCGCCAGCCCGGTGTCACTCGGCCCGGACACACGCGCCCGCG 60

RESULT 2

AAN90579/c
ID AAN90579 standard; cDNA; 309 BP.
XX
AC AAN90579;

DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 04-DEC-1989 (first entry)
XX
DE PDF9.3 cDNA insert.
XX

KW PDF9.3; human DF3 breast carcinoma-associated antigen epitope.
XX
OS Human MCF-7 breast carcinoma cells; 'ATCC HTB22'.
XX
PN W08907107-A.
XX
PD 10-AUG-1989.
XX

PF 29-JAN-1988; 88US-00149831.
PR 29-JAN-1988; 88US-00149831.
XX
PA (DANA-) DANA-FARBER CANCER.
XX
PI Kufe DW;
XX
DR WPI; 1989-248989/34.
DR P-PSDB; AAP91045, AAP91053, AAP91054, AAP90146.
XX
PT Recombinant polypeptide(s) - contains DF 3 breast carcinoma antigen
XX epitope and useful as assay reagents, and encoding DNA sequences.
PS Claim 1; Fig 4; 31pp; English.

XX The sequence encodes a carbohydrate-free polypeptide contg. DF3 breast
CC carcinoma antigen epitope. Useful as a competitive binding assay reagent
CC and improves diagnosis. The cDNA consists of nearly identical 60 BP
CC tandem repeats which are 85% GC-rich. See also AAP91045, AAP91046,
CC AAP91053, AAP91054. (Updated on 25-MAR-2003 to correct PA field.)
XX (Updated on 27-AUG-2003 to correct OS field.)
SQ Sequence 309 BP; 21 A; 92 C; 160 G; 36 T; 0 U; 0 Other;

Query Match
Best Local Similarity 100.0%; Score 60; DB 1; Length 309;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCTCCACCGCGCCCGCCAGCCCGGTGTCACTCGGCCCGGACACGCGCCCGCG 60
Db 266 GGCTCCACCGCGCCCGCCAGCCCGGTGTCACTCGGCCCGGACACGCGCCCGCG 207

RESULT 3

AD157712
ID AD157712 standard; cDNA; 1194 BP.
XX
AC AD157712;
XX

DT 22-APR-2004 (first entry)
XX

DE Human breast specific nucleic acid (BSNA) #83.
XX

KW Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
KW breast cancer; cytostatic.
XX

OS Homo sapiens.
XX

PN W02003106648-A2.
XX

PD 24-DEC-2003.
XX

PF 16-JUN-2003; 2003WO-US018934.
XX

PR 14-JUN-2002; 2002US-0389327P.
XX

PA (DIAD-) DIADEXUS INC.
XX

PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
XX
DR WPI; 2004-082185/08.
DR P-PSDB; AD157782.
XX

PT Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.
XX

PS Claim 1; SEQ ID NO 83; 370pp; English.
XX

CC The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the
CC hybridisation. The nucleic acids are useful for determining the presence
CC of a BSP in a sample which involves contacting the sample with suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein and
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence
CC of breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.
XX

SQ Sequence 1194 BP; 243 A; 438 C; 275 G; 238 T; 0 U; 0 Other;
Query Match
Best Local Similarity 100.0%; Score 60; DB 12; Length 1194;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCTCCACCGCGCCCGCCAGCCCGGTGTCACTCGGCCCGGACACGCGCCCGCG 60
Db 596 GGCTCCACCGCGCCCGCCAGCCCGGTGTCACTCGGCCCGGACACGCGCCCGCG 655

RESULT 4

AD157693
ID AD157693 standard; cDNA; 1378 BP.
XX
AC AD157693;

```
XX 22-APR-2004 (first entry)
DT
XX Human breast specific nucleic acid (BSNA) #64.
DE
XX Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
KM breast cancer; cytostatic.
XX
XX Homo sapiens.
OS
XX WO2003106648-A2.
PN
XX 24-DEC-2003.
PD
XX 16-JUN-2003; 2003WO-US018934.
PF
XX 14-JUN-2002; 2002US-0389327P.
PR
XX (DIAD-) DIADEXUS INC.
PA
XX Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
PI
XX WPI; 2004-082185/08.
DR
XX P-PSDB; ADI57765.
DR
XX
XX Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.
XX
XX Claim 1; SEQ ID NO 64; 370pp; English.
PS
XX
XX The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the presence
CC of a BSP in a sample which involves contacting the sample with suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein and
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.
XX
SQ Sequence 1378 BP; 295 A; 489 C; 316 G; 278 T; 0 U; 0 Other;
Query Match 100.0%; Score 60; DB 12; Length 1378;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTCCACCGCGCCCGCCAGCCGCGTGTCACTCGGCGCCGAGACACAGCGCGCGCCG 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 596 GGCTCCACCGCGCCCGCCAGCCGCGTGTCACTCGGCGCCGAGACACAGCGCGCGCCG 655
RESULT 5
ADO23180/c
ID ADO23180 standard; RNA; 1424 BP.
XX
XX ADO23180;
AC
XX 12-AUG-2004 (first entry)
DT
XX Antisense human MUC1 mucin glycoprotein RNA (coding sequence) SeqID 75.
DE
```

```
XX human; MUC1; mucin glycoprotein; cancer; chemotherapeutic; MUC1/BCD; ss;
KM antisense.
KM
XX Homo sapiens.
OS
XX WO2004044160-A2.
PN
XX 27-MAY-2004.
PD
XX 12-NOV-2003; 2003WO-US035848.
PF
XX 13-NOV-2002; 2002US-00293391.
PR
XX 29-MAY-2003; 2003US-00447839.
PR
XX (DAND ) DANA FARBER CANCER INST INC.
PA (ILEX-) ILEX PROD INC.
XX
XX Kufe DW, Kharbanda S, Weitman SD;
PI
XX WPI; 2004-420304/39.
DR
XX
XX Double-stranded RNA complex useful for inhibiting proliferation of cancer
PT cell expressing MUC1 mucin glycoprotein, comprises first and second RNA
PT sequences.
XX
XX Disclosure; SEQ ID NO 75; 112pp; English.
PS
XX
XX This invention relates to novel modulators of the human MUC1 mucin
CC glycoprotein for use in cancer therapeutics, where MUC1 is a protein that
CC acts to inhibit the apoptotic response to genotoxic stress caused by
CC chemotherapeutic agents. In particular, it refers to modulators of the
CC MUC1 extracellular domain (MUC1/BCD). The method refers to using double-
CC stranded RNA complexes as MUC1 interference RNA compositions such that
CC MUC1 expression is inhibited, which in turn inhibits cancer cell
CC proliferation. The present invention describes screening assays to
CC identify compounds that inhibit the binding of various MUC1 ligands such
CC as neuregulin 2, as well as agonists, antagonists and antibodies thereof.
CC Furthermore, it provides MUC1 antisense, small interfering RNA (siRNA)
CC and small molecules in combination with chemotherapeutic agents that are
CC useful in the field of cancer therapy. This polynucleotide sequence is
CC the antisense human MUC1 RNA of the invention.
XX
SQ Sequence 1424 BP; 309 A; 325 C; 495 G; 0 T; 295 U; 0 Other;
Query Match 100.0%; Score 60; DB 12; Length 1424;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTCCACCGCGCCCGCCAGCCGCGTGTCACTCGGCGCCCGAGACACAGCGCGCGCCG 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1041 GGCTCCACCGCGCCCGCCAGCCGCGTGTCACTCGGCGCCCGAGACACAGCGCGCGCCG 982
RESULT 6
ABL60159
ID ABL60159 standard; cDNA; 1428 BP.
XX
XX ABL60159;
AC
XX 22-JUL-2002 (first entry)
DT
XX Human MUC1 encoding cDNA SEQ ID NO 2.
DE
XX Human; mucin 1; MUC1; transmembrane protein; SNP; cancer; cytostatic;
KM single nucleotide polymorphism; haplotyping; genotyping; drug;
KM antiinflammatory; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH CDS 1..1428
FT /*tag= a
```

```
FT variation /product= "MUC1"
FT replace(1009,A)
FT /*tag= b
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP allelic variation results in Val substituted
FT by Met at position 337 of the MUC1 protein (ABB77476)"
XX PN WO200226765-A2.
XX PD 04-APR-2002.
XX PF 25-SEP-2001; 2001WO-US030151.
XX PR 28-SEP-2000; 2000US-0236113P.
XX PA (GENA-) GENAISSANCE PHARM INC.
XX PI Chew A, Koshiy B;
XX DR WPI; 2002-405042/43.
XX DR P-PSDB; ABB77476.
XX PT New genetic variants of mucin 1, Transmembrane gene, useful in studying
XX expression and function of protein encoded by the gene and for screening
XX drugs to treat diseases e.g. cancer.
XX PS Claim 23; Fig 2; 75pp; English.
XX CC The invention relates to a polynucleotide (ABL60158, ABL60159) encoding
XX mucin 1/MUC1 (ABB77476), Transmembrane isogene. The invention describes
XX novel genetic variants of the MUC1 gene. The invention is useful for
XX haplotyping/genotyping the MUC1 gene in an individual and identifying an
XX association between a trait and at least one of the haplotypes or
XX haplotype pairs of MUC1 gene. MUC1 is useful for studying the expression
XX and function of MUC1 and expressing MUC1 protein for use in screening for
XX candidate drugs to treat diseases related to MUC1 activity and in
XX studying the effect of the variation on the biological activity of MUC1
XX as well as on the binding affinity of candidate drugs targeting MUC1 for
XX the treatment of e.g. cancer. MUC1 is further used by the pharmaceutical
XX research scientist to validate MUC1 as a candidate target for and in
XX design of clinical trials of candidate drugs for, treating a specific
XX condition drugs or disease predicted to be associated with MUC1 activity.
XX CC MUC1 antibodies are useful in a variety of diagnostic and prognostic
XX formats and therapeutic methods
XX SQ Sequence 1428 BP; 296 A; 496 C; 327 G; 309 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 60; DB 6; Length 1428;
XX Best Local Similarity 100.0%; Pred. No. 1.4e-06;
XX Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GGCTCCACCGCCCCCGCCAGCCACGGGTGTCACTCGGCCCCGGAGACACGAGCCGCCCGG 60
Db 385 GGCTCCACCGCCCCCGCCAGCCACGGGTGTCACTCGGCCCCGGAGACACGAGCCGCCCGG 444
RESULT 7
ADO23125
ID ADO23125 standard; RNA; 1428 BP.
XX AC ADO23125;
XX DT 12-AUG-2004 (first entry)
XX DE Human MUC1 mucin glycoprotein RNA (coding sequence) SeqID 20.
XX KW human; MUC1; mucin glycoprotein; cancer; chemotherapeutic; MUC1/ECD; ss.
XX OS Homo sapiens.
XX PN WO200404160-A2.
XX PD 27-MAY-2004.
```

```
XX PF 12-NOV-2003; 2003WO-US035848.
XX PR 13-NOV-2002; 2002US-00293391.
XX PR 29-MAY-2003; 2003US-00447839.
XX PA (DAND ) DANA FARBER CANCER INST INC.
XX PA (ILEX-) ILEX PROD INC.
XX PI Kufe DW, Kharbanda S, Weltman SD;
XX DR WPI; 2004-420304/39.
XX PT Double-stranded RNA complex useful for inhibiting proliferation of cancer
XX cell expressing MUC1 mucin glycoprotein, comprises first and second RNA
XX sequences.
XX PS Claim 2; SEQ ID NO 20; 112pp; English.
XX CC This invention relates to novel modulators of the human MUC1 mucin
XX glycoprotein for use in cancer therapeutics, where MUC1 is a protein that
XX acts to inhibit the apoptotic response to genotoxic stress caused by
XX chemotherapeutic agents. In particular, it refers to modulators of the
XX MUC1 extracellular domain (MUC1/ECD). The method refers to using double-
XX stranded RNA complexes as MUC1 interference RNA compositions such that
XX MUC1 expression is inhibited, which in turn inhibits cancer cell
XX proliferation. The present invention describes screening assays to
XX identify compounds that inhibit the binding of various MUC1 ligands such
XX as neuiregulin 2, as well as agonists, antagonists and antibodies thereof.
XX Furthermore, it provides MUC1 antisense, small interfering RNA (siRNA)
XX and small molecules in combination with chemotherapeutic agents that are
XX useful in the field of cancer therapy. This polynucleotide sequence is
XX the human MUC1 RNA of the invention.
XX SQ Sequence 1428 BP; 296 A; 496 C; 327 G; 0 T; 309 U; 0 Other;
XX
XX Query Match 100.0%; Score 60; DB 12; Length 1428;
XX Best Local Similarity 93.3%; Pred. No. 1.4e-06;
XX Matches 56; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
OY 1 GGCTCCACCGCCCCCGCCAGCCACGGGTGTCACTCGGCCCCGGAGACACGAGCCGCCCGG 60
Db 385 GGCTCCACCGCCCCCGCCAGCCACGGGTGTCACTCGGCCCCGGAGACACGAGCCGCCCGG 444
RESULT 8
ADF32627
ID ADF32627 standard; DNA; 1457 BP.
XX AC ADF32627;
XX DT 26-FEB-2004 (first entry)
XX DE Plasmid JNM358 MUC-1 nucleotide sequence.
XX KW MUC-1 antigen; immune response; MUC-1; variable number of tandem repeat;
XX VNTR; repeat unit; tumour; metastasis; cytostatic; vaccine; gene therapy;
XX gene; ds.
XX OS Synthetic.
XX PN WO2003100060-A2.
XX PD 04-DEC-2003.
XX PF 23-MAY-2003; 2003WO-EP005594.
XX PR 24-MAY-2002; 2002GB-00012046.
XX PA (GLAX ) GLAXO GROUP LTD.
XX PI Burden N, Ellis JH, Hamblin PA;
```

DR WPI; 2004-042811/04.
XX
PT New nucleic acid molecule encoding a MUC-1 antigen, useful for preparing
a composition for treating or preventing tumors or metastases.
XX
PS Example; Fig 3; 66pp; English.
XX
CC The present invention describes a nucleic acid molecule which encodes a
MUC-1 antigen. The nucleic acid is capable of raising an immune response
in vivo, has reduced susceptibility to recombination than full-length MUC
-1 and comprises between 1 and 15 variable number of tandem repeats.
CC (VNTR) perfect repeat units. Also described: (1) a plasmid comprising the
CC DNA molecule; (2) a protein encoded by the nucleic acid; (3) a
CC pharmaceutical composition comprising the nucleic acid, plasmid or
CC protein and an excipient, diluent or carrier; and (4) a method of
CC treating or preventing tumors or metastases. A MUC1 antigen has
CC cytostatic activity, and can be used in vaccines, and in gene therapy.
CC The nucleic acid is useful for preparing a composition for treating or
CC preventing tumors or metastases. The present sequence is used in the
CC exemplification of the present invention.
XX
SQ Sequence 1457 BP; 302 A; 506 C; 334 G; 315 T; 0 U; 0 Other;
Query Match 100.0%; Score 60; DB 12; Length 1457;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTCCACCGCCCCCAGCCCGGTGTCACTTCGCCCCCGGACACGAGCCGCCCCG 60
DB 403 GGCTCCACCGCCCCCAGCCCGGTGTCACTTCGCCCCCGGACACGAGCCGCCCCG 462
RESULT 9
AAV48329
ID AAV48329 standard; cDNA; 1527 BP.
XX
AC AAV48329;
XX
DT 20-NOV-1998 (first entry)
XX
DE MiniMUC1 gene.
XX
KW 86; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW tumour; tumour-associated antigen.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1527
FT /*tag= a
FT /product= "MiniMUC1 protein"
XX
PN WO9837095-A2.
XX
PD 27-AUG-1998.
XX
PF 24-FEB-1998; 98WO-US003693.
XX
PR 24-FEB-1997; 97US-0038253P.
XX
PA (THER-) THERION BIOLOGICS CORP.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (DAND) DANA FARBER CANCER INST INC.
XX
PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX WPI; 1998-467492/40.
DR P-PSDB; AAW77233.
XX
XX New recombinant pox virus for tumour therapy - comprises DNA encoding an
immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
PT Example 1; Page 21-22; 42pp; English.

XX The immunogenic mini-MUC1 fragment was created from MUC1 tandem repeat
CC units for inclusion in a recombinant pox virus (RPV). The RPV was used in
CC a pharmaceutical composition also containing an immunomodulator to
CC generate MUC1 specific cytotoxic T-lymphocytes. The recombinant pox virus
CC therefore encodes an immunogenic MUC1 fragment that does not undergo
CC significant genetic deletion, thereby providing an unexpectedly stable
CC and immunogenic pox virus. They can be used to prevent or treat tumors
CC expressing MUC1 tumour-associated antigens
XX
SQ Sequence 1527 BP; 296 A; 573 C; 351 G; 307 T; 0 U; 0 Other;
Query Match 100.0%; Score 60; DB 2; Length 1527;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTCCACCGCCCCCAGCCCGGTGTCACTTCGCCCCCGGACACGAGCCGCCCCG 60
DB 226 GGCTCCACCGCCCCCAGCCCGGTGTCACTTCGCCCCCGGACACGAGCCGCCCCG 285
RESULT 10
ADK70370
ID ADK70370 standard; cDNA; 1614 BP.
XX
AC ADK70370;
XX
DT 06-MAY-2004 (first entry)
XX
DE Respiratory disease differentially expressed cDNA #106.
XX
KW de; gene; cytostatic; respiratory; antiasthmatic; gene therapy;
KW differential gene expression; respiratory disorder; lung cancer;
KW chronic obstructive pulmonary disease; emphysema; asthma.
XX
OS Homo sapiens.
XX
PN WO2003101283-A2.
XX
PD 11-DEC-2003.
XX
PF 02-JUN-2003; 2003WO-US017409.
XX
PR 04-JUN-2002; 2002US-0386005P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Rickett PK, Krasnow R;
XX
XX WPI; 2004-042945/04.
XX
DR New combination comprising cDNAs and proteins that are differentially
PT expressed in respiratory disorders, useful for diagnosing or treating
PT respiratory diseases e.g. lung cancer, chronic obstructive pulmonary
PT diseases or asthma.
XX
PS Claim 1; SEQ ID NO 106; 343pp; English.
XX
XX The invention relates to cDNA sequences that are differentially expressed
CC in respiratory disorders or their complements or encoded proteins. The
CC cDNAs and proteins are useful for diagnosing, treating or monitoring
CC treatment of a subject with a respiratory disease including lung cancer,
CC chronic obstructive pulmonary diseases, emphysema or asthma. The protein
CC is also useful for screening molecules or compounds to identify at least
CC one ligand which specifically binds the protein. It is also useful for
CC preparing and purifying a polyclonal or monoclonal antibody. This
CC sequence corresponds to a cDNA of the invention.
XX
SQ Sequence 1614 BP; 329 A; 556 C; 374 G; 353 T; 0 U; 2 Other;
Query Match 100.0%; Score 60; DB 12; Length 1614;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCGACGGGTGTACACCTCGGCCCCGGACACAGGCGGCCCGG 60
|||||
Db 434 GGCTCCACCGCCCCCGACGGGTGTACACCTCGGCCCCGGACACAGGCGGCCCGG 493

RESULT 11
AD157708
ID AD157708 standard; cDNA; 1630 BP.
XX
AC AD157708;
XX
DT 22-APR-2004 (first entry)
XX
DE Human breast specific nucleic acid (BSNA) #79.
XX
KM Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
KM breast cancer; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2003106648-A2.
XX
PD 24-DEC-2003.
XX
PF 16-JUN-2003; 2003WO-US018934.
XX
PR 14-JUN-2002; 2002US-0389327P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
XX
DR WPI; 2004-082185/08.
DR P-PSDB; AD157779.
XX
PT Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.
XX
PS Claim 1; SEQ ID NO 79; 370pp; English.
XX
CC The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the
CC hybridisation. The nucleic acids are useful for determining the presence
CC of a BSP in a sample which involves contacting the sample with suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein and
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.
XX
SQ Sequence 1630 BP; 322 A; 584 C; 405 G; 319 T; 0 U; 0 Other;

Query Match 100.0%; Score 60; DB 12; Length 1630;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCGACGGGTGTACACCTCGGCCCCGGACACAGGCGGCCCGG 60
|||||
Db 596 GGCTCCACCGCCCCCGACGGGTGTACACCTCGGCCCCGGACACAGGCGGCCCGG 655

RESULT 12
AD157689
ID AD157689 standard; cDNA; 1634 BP.
XX
AC AD157689;
XX
DT 22-APR-2004 (first entry)
XX
DE Human breast specific nucleic acid (BSNA) #60.
XX
KM Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
KM breast cancer; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2003106648-A2.
XX
PD 24-DEC-2003.
XX
PF 16-JUN-2003; 2003WO-US018934.
XX
PR 14-JUN-2002; 2002US-0389327P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
XX
DR WPI; 2004-082185/08.
DR P-PSDB; AD157761.
XX
PT Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.
XX
PS Claim 1; SEQ ID NO 60; 370pp; English.

XX
CC The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the
CC hybridisation. The nucleic acids are useful for determining the presence
CC of a BSP in a sample which involves contacting the sample with suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein and
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.
XX
SQ Sequence 1634 BP; 349 A; 541 C; 405 G; 337 T; 0 U; 2 Other;

Query Match 100.0%; Score 60; DB 12; Length 1634;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCGACGGGTGTACACCTCGGCCCCGGACACAGGCGGCCCGG 60
|||||
Db 596 GGCTCCACCGCCCCCGACGGGTGTACACCTCGGCCCCGGACACAGGCGGCCCGG 655

RESULT 13
AD157686

```
ID ADI57686 standard; cDNA; 1712 BP.
XX
XX ADI57686;
AC
XX
DT 22-APR-2004 (first entry)
XX
XX Human breast specific nucleic acid (BSNA) #57.
DE
XX Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
KM breast cancer; cytostatic.
XX
XX Homo sapiens.
OS
XX WO2003106648-A2.
PN
XX 24-DEC-2003.
PD
XX 16-JUN-2003; 2003WO-US018934.
PF
XX 14-JUN-2002; 2002US-0389327P.
PR
XX (DIAD-) DIADEXUS INC.
PA
XX Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
PI
XX WPI; 2004-082185/08.
DR
XX P-PSDB; ADI57758.
DR
XX
XX Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.
XX
XX Claim 1; SEQ ID NO 57; 370pp; English.
XX
XX The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the
CC hybridisation. The nucleic acids are useful for determining the presence
CC of a BSP in a sample which involves contacting the sample with suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein in a
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.
XX
XX Sequence 1712 BP; 327 A; 604 C; 435 G; 346 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 60; DB 12; Length 1712;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTCCACCGCCCCCGACCGGTGTCACTCGGCCCGGACACACGCGCGCCCG 60
Db 596 GGCTCCACCGCCCCCGACCGGTGTCACTCGGCCCGGACACACGCGCGCCCG 655
RESULT 14
ADI57669 standard; cDNA; 1738 BP.
XX
XX ADI57669;
AC
XX
```

```
DT 22-APR-2004 (first entry)
XX
XX Human breast specific nucleic acid (BSNA) #40.
DE
XX Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
KM breast cancer; cytostatic.
XX
XX Homo sapiens.
OS
XX WO2003106648-A2.
PN
XX 24-DEC-2003.
PD
XX 16-JUN-2003; 2003WO-US018934.
PF
XX 14-JUN-2002; 2002US-0389327P.
PR
XX (DIAD-) DIADEXUS INC.
PA
XX Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
PI
XX WPI; 2004-082185/08.
DR
XX P-PSDB; ADI57743.
DR
XX
XX Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.
XX
XX Claim 1; SEQ ID NO 40; 370pp; English.
XX
XX The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the
CC hybridisation. The nucleic acids are useful for determining the presence
CC of a BSP in a sample which involves contacting the sample with suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein in a
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.
XX
XX Sequence 1738 BP; 342 A; 642 C; 399 G; 355 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 60; DB 12; Length 1738;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTCCACCGCCCCCGACCGGTGTCACTCGGCCCGGACACACGCGCGCCCG 60
Db 458 GGCTCCACCGCCCCCGACCGGTGTCACTCGGCCCGGACACACGCGCGCCCG 517
RESULT 15
ADI57673 standard; cDNA; 1755 BP.
XX
XX ADI57673;
AC
XX 22-APR-2004 (first entry)
DT
XX Human breast specific nucleic acid (BSNA) #44.
DE
XX
```


KW Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
KW breast cancer; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2003106648-A2.
XX
PD 24-DEC-2003.
XX
PF 16-JUN-2003; 2003WO-US018934.
XX
PR 14-JUN-2002; 2002US-0389327P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
XX
DR WPI; 2004-082185/08.
DR P-PSDB; ADI57746.
XX
PT Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.
XX
PS Claim 1; SEQ ID NO 44; 370bp; English.
XX
CC The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the
CC hybridisation. The nucleic acids are useful for determining the presence
CC of a BSP in a sample which involves contacting the sample with suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein and
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.
XX
SQ Sequence 1755 BP; 348 A; 595 C; 432 G; 378 T; 0 U; 2 Other;

Query Match 100.0%; Score 60; DB 12; Length 1755;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCGCCAGCCAGGTGTCACTCGGCCCGGACACCAAGCCGGCCCCG 60
|||
Db 569 GGCTCCACCGCCCCCGCCAGCCAGGTGTCACTCGGCCCGGACACCAAGCCGGCCCCG 628

Search completed: January 15, 2005, 20:36:05
Job time : 171.5 secs

QY 1 GGCTCCACCGCCCCCGACCGCGGTGTACCTCGCCCCCGGACACCAAGCGGCCCCG 60
|||||
Db 316 GGCTCCACCGCCCCCGACCGCGGTGTACCTCGCCCCCGGACACCAAGCGGCCCCG 257
|||||

RESULT 2
BM791359 604 bp mRNA linear EST 05-MAR-2002
LOCUS K-EST0071342 S21SNUS20 Homo sapiens cDNA clone S21SNUS20-14-A06 5',
DEFINITION mRNA sequence.
BM791359
ACCESSION BM791359.1 GI:19139591
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 604)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
TITLE Unpublished (2002)
JOURNAL
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 14 row: A column: 06
High quality sequence stop: 604.
Location/Qualifiers
1. .604
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S21SNUS20-14-A06"
/sex="F"
/tissue_type="Stomach"
/cell_type="floating aggregates"
/cell_line="SNU-520"
/lab_host="Top10F"
/clone_lib="S21SNUS20"
/note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dT-selected mRNA by
priming with dT-tailed vector. The dT-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN

Query Match 100.0%; Score 60; DB 4; Length 604;
Best Local Similarity 100.0%; Pred. No. 5.9e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCGACCGCGGTGTACCTCGCCCCCGGACACCAAGCGGCCCCG 60
|||||
Db 106 GGCTCCACCGCCCCCGACCGCGGTGTACCTCGCCCCCGGACACCAAGCGGCCCCG 165
|||||

RESULT 3
BU542454

LOCUS BU542454 877 bp mRNA linear EST 13-SEP-2002
DEFINITION AGENCOURT 10322173 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574322
5', mRNA sequence.
BU542454
ACCESSION BU542454.1 GI:22852937
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 877)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LCM2769 row: h column: 02
High quality sequence stop: 760.
Location/Qualifiers
1. .877
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6574322"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 60; DB 5; Length 877;
Best Local Similarity 100.0%; Pred. No. 5.7e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCGACCGCGGTGTACCTCGCCCCCGGACACCAAGCGGCCCCG 60
|||||
Db 35 GGCTCCACCGCCCCCGACCGCGGTGTACCTCGCCCCCGGACACCAAGCGGCCCCG 94
|||||

RESULT 4
CA489836 959 bp mRNA linear EST 14-NOV-2002
LOCUS AGENCOURT 10810668 MAPCL Homo sapiens cDNA clone IMAGE:6722324 5',
DEFINITION mRNA sequence.
CA489836
ACCESSION CA489836.1 GI:24952627
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 959)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Kristi A. Egland, Ira Pastan
cDNA Library Preparation: Invitrogen Corp

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM14284 row: n column: 20
High quality sequence start: 31
High quality sequence stop: 446.

FEATURES

source

1..959

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6722324"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231,
hTERT-HME1, LNCaP"
/lab_host="EMDH10B"
/clone_lib="MAPCL"
/note="Vector: pCMV-SPORT6; Site_1: EcoRV; Site_2: Not I;
Subtracted with brain, liver, lung, kidney and muscle.
Directionally cloned. Priming method: oligo-dT. Average
insert size: 1800 bp. Library amplification: 26,000 fold.
Kristi A. Egland, James J. Vincent, Robert Strausberg,
Bungkook Lee & Ira Pastan: Discovery of new breast
cancer genes encoding membrane and secreted proteins.
Manuscript submitted."

ORIGIN

Query Match 100.0%; Score 60; DB 6; Length 959;
Best Local Similarity 100.0%; Pred. No. 5.7e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCAGCCCGGTGTCACTCGGCCCGGACACCAAGCCGCCCG 60
|||||
Db 349 GGCTCCACCGCCCCCAGCCCGGTGTCACTCGGCCCGGACACCAAGCCGCCCG 408

RESULT 5

LOCUS

BU148487 1113 bp mRNA linear EST 03-SEP-2002
AGENCOURT_8670479 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6380642

DEFINITION

5', mRNA sequence.

ACCESSION

VERSION

BU148487.1 GI:22662019

KEYWORDS

EST.

SOURCE

ORGANISM

Homo sapiens (human)
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
1 (bases 1 to 1113)

NIH-MGC http://mgs.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov

COMMENT

Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov

Plate: LLM2569 row: j column: 03
High quality sequence stop: 235.

FEATURES

source

1..1113

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6380642"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"

/note="Organ: prostate; Vector: pOT87; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACGAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 60; DB 5; Length 1113;
Best Local Similarity 100.0%; Pred. No. 5.6e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCAGCCCGGTGTCACTCGGCCCGGACACCAAGCCGCCCG 60
|||||
Db 90 GGCTCCACCGCCCCCAGCCCGGTGTCACTCGGCCCGGACACCAAGCCGCCCG 149

RESULT 6

LOCUS

BU542996 1130 bp mRNA linear EST 13-SEP-2002
AGENCOURT_10338707 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574979

DEFINITION

5', mRNA sequence.

ACCESSION

BU542996

VERSION

BU542996.1 GI:22853479

KEYWORDS

EST.

SOURCE

Homo sapiens (human)
Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
1 (bases 1 to 1130)

NIH-MGC http://mgs.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov

COMMENT

Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov

Plate: LLM2771 row: c column: 11
High quality sequence start: 27

High quality sequence stop: 246.

FEATURES

source

1..1130

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6574979"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOT87; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACGAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 60; DB 5; Length 1130;
Best Local Similarity 100.0%; Pred. No. 5.6e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCAGCCCGGTGTCACTCGGCCCGGACACCAAGCCGCCCG 60
|||||
Db 69 GGCTCCACCGCCCCCAGCCCGGTGTCACTCGGCCCGGACACCAAGCCGCCCG 128

RESULT 7
BQ936898
LOCUS
DEFINITION BQ936898 1234 bp mRNA linear EST 21-AUG-2002
AGENCOURT_8919110 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6383855
5', mRNA sequence.
BQ936898
ACCESSION BQ936898.1 GI:22352281
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 1234)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LNCM2577 row: 0 column: 24
High quality sequence stop: 245.
Location/Qualifiers
1..1234
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6383855"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACGAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

FEATURES
source
Query Match 100.0%; Score 60; DB 5; Length 1234;
Best Local Similarity 100.0%; Pred. No. 5.6e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCAGCCACGGGTGTCACCTCGGCCCCGGACACAGCGCCCCG 60
90 GGCTCCACCGCCCCCAGCCACGGGTGTCACCTCGGCCCCGGACACAGCGCCCCG 149

Db

RESULT 8
BQ943554
LOCUS
DEFINITION BQ943554 1268 bp mRNA linear EST 21-AUG-2002
AGENCOURT_8777815 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6384308
5', mRNA sequence.
BQ943554
ACCESSION BQ943554.1 GI:22359032
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 1268)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
COMMENT Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LNCM2579 row: b column: 21
High quality sequence stop: 177.
Location/Qualifiers
1..1268
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6384308"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACGAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

ORIGIN
Query Match 100.0%; Score 60; DB 5; Length 1268;
Best Local Similarity 100.0%; Pred. No. 5.5e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCAGCCCAAGGTGTACCTCGGCCCCGGACACAGCGCCCCG 60
90 GGCTCCACCGCCCCCAGCCCAAGGTGTACCTCGGCCCCGGACACAGCGCCCCG 149

Db

RESULT 9
BQ920055
LOCUS
DEFINITION BQ920055 1343 bp mRNA linear EST 20-AUG-2002
AGENCOURT_10031674 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6481609
5', mRNA sequence.
BQ920055
ACCESSION BQ920055.1 GI:22334753
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 1343)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LNCM2663 row: e column: 02
High quality sequence start: 56
High quality sequence stop: 237.
Location/Qualifiers
1..1343
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6481609"

FEATURES
source

/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACGAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 60; DB 5; Length 1343;
Best Local Similarity 100.0%; Pred. No. 5.5e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCAGCCGAGGTGTCACTCGGCCCGGACACAGCGCCGCCCG 60
|||||
Db 132 GGCTCCACCGCCCCCAGCCGAGGTGTCACTCGGCCCGGACACAGCGCCGCCCG 191

RESULT 10
BU152566 1349 bp mRNA linear EST 03-SEP-2002
LOCUS AGENCOURT_8670683 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6380559
DEFINITION 5', mRNA sequence.
ACCESSION BU152566
VERSION BU152566.1 GI:22666098
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE 1 (bases 1 to 1349)
JOURNAL NIH-MGC http://mgc.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLCM2569 row: f column: 16
High quality sequence stop: 291.

FEATURES

source

1.1349
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6380559"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACGAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 60; DB 5; Length 1349;
Best Local Similarity 100.0%; Pred. No. 5.5e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCAGCCGAGGTGTCACTCGGCCCGGACACAGCGCCGCCCG 60

Db 19 GGCTCCACCGCCCCCAGCCGAGGTGTCACTCGGCCCGGACACAGCGCCGCCCG 78
|||||

RESULT 11
BU542790 1420 bp mRNA linear EST 13-SEP-2002
LOCUS AGENCOURT_10334841 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574725
DEFINITION 5', mRNA sequence.
ACCESSION BU542790
VERSION BU542790.1 GI:22853273
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE 1 (bases 1 to 1420)
JOURNAL NIH-MGC http://mgc.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLCM2770 row: h column: 21
High quality sequence stop: 288.

FEATURES

source

1.1420
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6574725"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACGAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 60; DB 5; Length 1420;
Best Local Similarity 100.0%; Pred. No. 5.5e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCAGCCGAGGTGTCACTCGGCCCGGACACAGCGCCGCCCG 60
|||||
Db 19 GGCTCCACCGCCCCCAGCCGAGGTGTCACTCGGCCCGGACACAGCGCCGCCCG 78

RESULT 12
BU543309 1531 bp mRNA linear EST 13-SEP-2002
LOCUS AGENCOURT_10327072 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6575349
DEFINITION 5', mRNA sequence.
ACCESSION BU543309
VERSION BU543309.1 GI:22853792
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE 1 (bases 1 to 1531)
JOURNAL NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLCM2772 row: b column: 21
High quality sequence stop: 166.
Location/Qualifiers
1. .1531
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6575349"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN
Query Match 97.3%; Score 58.4; DB 5; Length 1531;
Best Local Similarity 98.3%; Pred. No. 1.4e-05;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCAGCCACGGTGTCACTCGGCCCGGACACACGCGGCCCG 60
|||||
30 GGCTCCACCGCCCCCAGCCACGGTGTCACTCGGCCCGGACACACGCGGCCCG 89

Db 30 GGCTCCACCGCCCCCAGCCACGGTGTCACTCGGCCCGGACACACGCGGCCCG 89

RESULT 13
BQ935496 1262 bp mRNA linear EST 21-AUG-2002
LOCUS AGENCOURT_8776475 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6382088
DEFINITION 5', mRNA sequence.
ACCESSION BQ935496
VERSION BQ935496.1 GI:22350879
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1262)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLCM2573 row: f column: 09
High quality sequence start: 46
High quality sequence stop: 157.
Location/Qualifiers
1. .1262
/organism="Homo sapiens"
/mol_type="mRNA"

FEATURES
source

/db_xref="taxon:9606"
/clone="IMAGE:6382088"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN
Query Match 94.7%; Score 56.8; DB 5; Length 1262;
Best Local Similarity 96.7%; Pred. No. 3.8e-05;
Matches 58; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCAGCCACGGTGTCACTCGGCCCGGACACACGCGGCCCG 60
|||||
56 GGCTCCACCGCCCCCAGCCACGGTGTCACTCGGCCCGGACACACGCGGCCCG 115

Db 56 GGCTCCACCGCCCCCAGCCACGGTGTCACTCGGCCCGGACACACGCGGCCCG 115

RESULT 14
BG774910 981 bp mRNA linear EST 15-MAY-2001
LOCUS BG774910
DEFINITION mRNA sequence.
ACCESSION BG774910
VERSION BG774910.1 GI:14045227
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 981)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLCM1612 row: o column: 07
High quality sequence stop: 874.
Location/Qualifiers
1. .981
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4761054"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN
Query Match 87.0%; Score 52.2; DB 4; Length 981;
Best Local Similarity 94.7%; Pred. No. 0.00062;
Matches 54; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCGCCCGCCAGCCCGGTGTACCTCGGCGCCGGACACAGCGCGGCC 57
 |||||
 Db 21 GGCTCCACCGCGCCCGCCAGCCCGGTGTACCTCGGCGCCGGACACAGCGCGGCC 77
 |||||

RESULT 15
 BQ923149 1536 bp mRNA linear EST 20-AUG-2002
 LOCUS BQ923149
 DEFINITION AGENCOURT 8929207 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6484568
 5' mRNA sequence.

ACCESSION BQ923149
 VERSION BQ923149.1 GI:22338180
 KEYWORDS EST.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1536)
 NIH-MGC http://mgs.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: DCTD/DTF

CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLCM2670 row: p column: 09
 High quality sequence stop: 287.
 Location/Qualifiers

FEATURES

source 1..1536
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6484568"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NIH_MGC_40"
 /note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
 Site_2: EcoRI; cDNA made by oligo-dT priming.
 directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGACGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 87.0%; Score 52.2; DB 5; Length 1536;
 Best Local Similarity 94.7%; Pred. No. 0.00059;
 Matches 54; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 GGCTCCACCGCGCCCGCCAGCCCGGTGTACCTCGGCGCCGGACACAGCGCGGCC 57
 |||||
 Db 18 GGCTCCACCGCGCCCGCCAGCCCGGTGTACCTCGGCGCCGGACACAGCGCGGCC 74
 |||||

Search completed: January 16, 2005, 02:55:18
 Job time : 1555.6 secs

This Page Blank (uspto)